

Synthesis of Chemical Tools for Targeted Epigenetic Modifications



Ha Phuong Nguyen

Institute of Structural and Molecular Biology
University College London

This dissertation is submitted for the degree of
Doctor of Philosophy

February 2019

Declaration

I, Ha Phuong Nguyen, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Ha Phuong Nguyen
February 2019

Abstract

DNA methylation is an important epigenetic modification that regulates gene expression, thereby playing a crucial role in biological processes including cellular development but also disease progression. Abnormal methylation patterns such as gene promoter hypermethylation are the most common molecular lesions found in cancer. Since DNA methylation is a reversible process, DNA methyltransferases (DNMTs) – the enzymes responsible for the maintenance of methylation patterns - are regarded as very attractive therapeutic targets.

To date, several DNMT inhibitors have been developed to reverse DNA hypermethylation. The clinically approved azanucleosides 5-azacytidine and 5-aza-2'-deoxycytidine have high demethylating potency, but their use has remained limited due to their nonspecific activity and cytotoxic side-effects. The lack of techniques for targeted tissue-specific demethylation not only limits therapeutic potential but also restricts our understanding of the functional significance of DNA methylation in a number of biological processes, including cellular development and regeneration.

The aim of this thesis was to develop small molecule tools that could achieve spatiotemporally targeted demethylation through light-based activation. We have used a photocaging strategy to synthesise six novel photocaged azanucleoside analogues. These analogues remained biologically inert until light illumination induced the release of active azanucleosides with moderate-to-high rates and efficiencies. It was found that the caging position strongly influenced the photochemical properties and aqueous stabilities of the photocaged analogues. The analogue with the fastest uncaging rate and best aqueous stability 3'-DEACMOC-dAC was evaluated for its biological performance in a number of cancer cell lines. Global DNA methylation analysis showed that the analogue achieved light-dependent demethylation at lower concentration treatments. However, at higher concentrations unexpected activity was observed in the absence of light, prompting further investigation using mass spectrometry analysis.

Impact statement

DNA methylation plays an essential part in maintaining and regulating key cellular events. The ability to understand the functional changes in DNA methylation and finely manipulate these changes (e.g. in order to prevent the development of cancer and other diseases) can serve as a significant stepping stone in biomedical and cancer research. Although a range of biological tools (e.g. CRISPR-Cas9,⁸³ TALE-TET1 fusion proteins⁷⁷) have been developed recently to study and induce DNA methylation, there have been fewer chemical tools that could achieve similar effects.

This work presents the first example of a light-gated small molecule approach to regulate DNA methylation within the cells. It was demonstrated that the activity of decitabine, inhibitor of DNA methyltransferase and a key epigenetic drug, can be masked by its synthetic modification (also known as “photocaging”) with a phototag, a photochemically removable functional group.

The impact of this work can be beneficial primarily in academia, and if further improved, potentially outside of academia too. The work provides novel insights on the chemistry of azanucleosides (decitabine, azacytidine) and how the photocaging and linker strategy can affect photochemical performance and the compound’s aqueous stability profile. The work also serves as a useful groundwork for the setup of photochemical, biological and stability characterisation of current and future photocaged azanucleosides.

In their current form, photocaged azanucleosides require further improvement in their stability profiles in order to be used as research tools to study epigenetic modifications. For clinical applications and therapeutic impact, the photocaged compounds could be further tailored to absorb at even higher wavelengths ($\lambda > 500$ nm) in order to be biocompatible with thicker tissues and organs. Once it is achieved, this strategy can potentially open up new avenues in basic and clinical research for the synthesis of other photo-controlled epigenetic molecules or future improved versions of other epigenetic drugs.

Acknowledgements

First and foremost I would like to thank my supervisors Prof. Stefan Howorka and Prof. Stephan Beck for giving me the opportunity to work on this exciting interdisciplinary project. I am extremely grateful for their guidance, constant enthusiasm, optimism and support along the way.

I am very thankful to Prof. Jim Anderson for his help with organic chemistry, Dr Mikie Kukwikila for laying the groundwork for this project and Dr Robin Bofinger for his help with HPLC and chemistry in general. I would also like to thank Dr Abil Aliev and Dr Kersti Karu for their help with NMR and mass spectrometry.

Special thanks goes to the UCL Medical Genomics group for their kindness and patience with a chemist like myself. In particular I would like to thank Dr Sabrina Stewart for her help in the early years of the project and for teaching me basic cell culture techniques. My sincere gratitude goes to Dr Andy Feber for patiently overseeing the progress of my biological experiments along the way and providing useful insights. Thanks to Dr Anna Köferle and Dr Ron Finn for their help in running and troubleshooting Western blots. For their support with the triple quadrupole mass spectrometry, I am deeply grateful to Dr Andrew Weston and Dr Emmanuel Samuel.

I also wish to acknowledge our collaborators, Prof. Shankar Balasubramanian and Dr Shiqing Mao for their help with the LC-MS analysis. Their contribution was invaluable to the progress of this project.

Thank you to Prof. Carolyn Moores and Dr Jadranka Butorac for coordinating the PhD course and for their administrative support. I would also like to thank the BBSRC and UCL for funding my project.

I wish to thank all past and present members of the Howorka group and office 330 for making my time at UCL a truly memorable one – thank you Gen, Jo, Patrick, Conor, Daniel, Oni, Gurpreet, Yasmin, Yaomin and Yongzheng. A special thanks goes to

Matthew, Diana, Roberta, Francesco and Lien for the lively and entertaining lunches and valuable memories.

Finally, I would like to thank my parents for their unconditional support and for providing me with all the means to receive a good education and a life full of opportunities. Thanks to my (not so) little sister Hajni, for increasingly becoming my big sister and taking good care of me especially towards the end of my PhD. And Siddharth – this journey would not have been possible without your love, support and encouragement. Thank you for everything.

Table of contents

Declaration	iv
Abstract	vi
Impact statement	viii
Acknowledgements	xi
List of Figures	xvii
List of Tables	xxi
List of Schemes	xxiv
Acronyms	xxvi
1 Introduction	1
1.1 Epigenetics	1
1.2 DNA methylation	4
1.2.1 DNA methylation in mammals	4
1.2.2 DNA methyltransferases	4
1.2.3 TET enzymes	6
1.2.4 Disruption of methylation pattern in disease	8
1.3 Strategies to target DNA methylation	10
1.3.1 Biological editing tools	11
1.3.2 Chemical tools I: non-nucleoside inhibitors	13
1.3.3 Chemical tools IIa: nucleoside inhibitors	14
1.3.4 Chemical tools IIb: nucleoside inhibitor analogues	22
1.3.5 Need for controlled DNA methylation	24

1.4	Tuning bioactivity using light: Photocaging	26
1.4.1	Principles of photocaging	26
1.4.2	Phototags	27
1.4.3	Photocaging nucleosides	34
1.5	Aims of the project	40
2	Synthesis of Photocaged Azanucleosides	41
2.1	Introduction	41
2.2	First generation photocaged azanucleosides	41
2.2.1	Design and synthesis	42
2.2.2	Photochemical characterisation	45
2.3	Second generation photocaged azanucleosides	49
2.3.1	Design and synthesis	49
2.3.2	Photochemical characterisation	52
2.3.3	Stability studies	53
2.4	Third generation photocaged azanucleosides	60
2.4.1	Pre-assessment of carbonate linker stability	60
2.4.2	Synthesis of 5'-caged analogue	61
2.4.3	Synthesis of 3'-caged analogue	68
2.4.4	Photochemical characterisation	69
2.4.5	Stability studies	72
2.5	Summary	74
3	Bioactivity of Photocaged Azanucleosides	75
3.1	Introduction	75
3.2	DNA methylation assays in literature	76
3.2.1	Importance of DNA methylation analysis techniques	76
3.2.2	Locus-specific DNA methylation analysis techniques	76
3.2.3	Global methylation analysis techniques	78
3.3	Selecting a suitable DNA methylation assay	80
3.3.1	Trial: ELISA assays	81
3.3.2	Trial: LC-MS/MS analysis	83
3.4	DNA methylation analysis <i>via</i> LC-MS/MS	85
3.4.1	Optimisation of treatment conditions	85
3.4.2	Measurement of DNA methylation changes	86
3.4.3	Summary of LC-MS/MS methylation results	90
3.5	Western blot analysis	92

3.5.1	Western blot analysis of Saos-2 cell extracts	93
3.5.2	Western blot analysis of T24 cell extracts	94
3.5.3	Summary of western blot results	95
3.6	Cell viability assays	96
3.6.1	Aims	96
3.6.2	Effect under low photocaged dAC concentrations	96
3.6.3	Effect under high photocaged dAC concentrations	102
3.6.4	Summary of cell viability assays	104
3.7	Preliminary cell assays on previous analogues	104
3.7.1	Aim of preliminary experiments	104
3.7.2	Cell viability assays on N-DEACMOC-dAC	105
3.7.3	450K Methylation Array	106
3.8	Summary of cell biology experiments	108
3.8.1	Limitations and suggestions for further improvements	109
4	Stability of Photocaged Azanucleosides	111
4.1	Introduction	111
4.2	Literature on carbonate link stability	112
4.3	Hypothesis on the sources of instability and aims	113
4.4	Stability studies <i>via</i> HPLC	115
4.5	Stability studies <i>via</i> mass spectrometry	116
4.5.1	Triple quadrupole mass spectrometer	116
4.5.2	MS stability studies in water	117
4.5.3	MS stability studies in HEPES buffer pH 7.2	120
4.6	Summary of stability studies	122
5	Non-hydrolysable Azanucleosides	127
5.1	Introduction	127
5.2	Synthesis of non-hydrolysable Azanucleosides	128
5.3	Stability analysis of non-hydrolysable Azanucleosides	130
5.4	Suggested future work	132
6	Conclusion	133
6.1	Summary and aim of the work	133
6.2	Impact and significance of the work	134
6.2.1	Significance of the work	134
6.2.2	Related work described in the field	135

6.3	Limitations of the work	138
6.3.1	Limitation 1: Hydrolytic Instability	138
6.3.2	Limitation 2: The setup of cell biological experiments	144
6.3.3	Limitation 3: Stability Experiments	145
6.4	Suggested future work	148
Appendix A Experimental		151
A.1	General experimental	151
A.1.1	Chemicals and solvents	151
A.1.2	Chromatography	151
A.1.3	Spectroscopy	151
A.1.4	HPLC	153
A.1.5	Photouncaging experiments	153
A.1.6	Quantum yield measurements	153
A.1.7	Quantum yield calculations	155
A.1.8	Rate calculations	155
A.2	Characterisations	157
A.3	Cell biology experiments	181
A.3.1	Cell lines and reagents	181
A.3.2	Cell culture	181
A.3.3	Photocaged azanucleoside treatments	182
A.3.4	Global methylation analysis <i>via</i> LC-MS/MS	185
A.3.5	Statistical analysis of methylation values	185
A.3.6	Methylation analysis <i>via</i> the Infinium 450K BeadChip	185
A.3.7	Reagents for western blot analysis	186
A.3.8	Western blot analysis	186
A.3.9	Cell viability assays	192
A.3.10	ELISA assays	194
A.4	Mass spectrometry analysis	197
A.4.1	Triple quadrupole mass spectrometer	197
A.4.2	MRM optimisation	197
A.4.3	MRM LC-MS calibration of analytes	199
A.4.4	Sample preparation and stability experiments	200
A.4.5	Quantification	200
A.4.6	Summary of measured stability profiles	200
Appendix B Publications		203

List of Figures

1.1	Waddington's epigenetic landscape	2
1.2	Structure of the chromosome and summary of epigenetic modifications	3
1.3	Catalytic mechanism of cytosine methylation by DNMT1	5
1.4	Catalytic mechanism of oxidation of 5mC to 5hmC by TET enzymes .	6
1.5	Dynamic modifications of cytosine analogues	7
1.6	DNA methylation patterns in normal and cancer cells	9
1.7	Models of biological editing tools	11
1.8	Uptake and incorporation of azanucleosides	15
1.9	Mechanism of inhibition by azanucleosides	16
1.10	Summary of nucleoside caging strategies	34
1.11	Aims of the project	40
2.1	UV-Vis absorption profiles of first generation photocaged azanucleosides	45
2.2	Photouncaging profile of bis-NPEOC-AC 59 at $\lambda = 254$ nm	47
2.3	Photouncaging profiles of first generation photocaged dAC analogues 60 and 64	48
2.4	Photochemical characterisation of N-DEACMOC-dAC 78	52
2.5	HPLC traces for photouncaging of N-DEACMOC-dAC 78 at $\lambda = 365$ nm	53
2.6	HPLC stability studies of N-DEACMOC-dAC 78 at 37°C	54
2.7	MS profile of N-DEACMOC-dAC degradation product	56
2.8	HPLC traces of purified 5'-DEACMOC-dAC 'crude'	65
2.9	NMR analysis of 'crude' mixture from the synthesis of 5'-DEACMOC dAC 84	66
2.10	UV-Vis absorption profiles of DEACM-caged analogues 78 , 84 and 93	69
2.11	Photouncaging profiles of third generation photocaged azanucleosides 84 and 93	70
2.12	Photodeprotection of 1:1 5'- and 5'caged dAC mixture	71

2.13	Stability profiles of dAC, DEACM and analogues 93 and 84 at 37°C	73
3.1	Light tunable bioactivity of photocaged azanucleosides	76
3.2	Results of ELISA-based methylation assay trials	82
3.3	Results of LC-MS/MS methylation assay trial	84
3.4	Cell line screening in response to dAC treatment	86
3.5	LC-MS methylation analysis of SaOS-2 cells	87
3.6	LC-MS methylation analysis of T24 cells	89
3.7	LC-MS methylation analysis of HEK293T cells	91
3.8	Western blots of Saos-2 cell extracts	93
3.9	Western blots of T24 cell extracts	94
3.10	Cell viability assays of Saos-2 cells at low concentration treatments	97
3.11	Cell viability assays of T24 cells at low concentration treatments	98
3.12	Cell viability assays of HEK293T cells at low concentration treatments	99
3.13	Cell viability assays of Saos-2 cells at high concentration treatments	103
3.14	Cell viability assays of Saos2 cells in response to N-DEACMOC-dAC I	106
3.15	Cell viability assays of Saos2 cells in response to N-DEACMOC-dAC II	107
3.16	450K Array results	107
4.1	HPLC stability analysis of dAC and photocaged dAC at 37°C	115
4.2	Schematic diagram of a triple quadrupole mass spectrometer	116
4.3	RT-MS Stability profiles of photocaged analogues in water	117
4.4	37°C-MS Stability profiles of photocaged analogues in water	118
4.5	Quantification of analytes during 3'-CD incubation at 37°C in water	119
4.6	MS Stability profiles of photocaged analogues in HEPES pH 7.2	121
4.7	Comparison of calculated and experimentally measured degradation curves of 3'-caged dAC	123
5.1	HPLC Stability analysis of DMT-dAC, 97	130
5.2	Purification of DMT-dAC, 97	131
5.3	HPLC Stability analysis of twice purified DMT-dAC, 97	131
6.1	Stability profiles on carbonate linked 5-iodo-2'-deoxyuridine in various physiological media	141
A.1	Linear Regression plots of Photouncaging Decays	156
A.2	Electrophoretic transfer setup	187
A.3	Relative DNMT1 band intensities in 3'-CD treated Saos-2 cells	190

A.4	Relative DNMT1 band intensities in 5'-CD treated Saos-2 cells	191
A.5	Relative DNMT1 band intensities in 3'-CD treated T24 cells	191
A.6	Relative DNMT1 band intensities in 5'-CD treated T24 cells	191
A.7	MRM LC-MS calibration of dAC, DEACM and 3'-CD	199

List of Tables

1.1	Frequently hypermethylated genes in sporadic cancers	10
2.1	UV-Vis absorption properties of first generation photocaged azanucleosides	46
2.2	Photochemical properties of first generation photocaged azanucleosides	49
2.3	Stability of N-DEACMOC-dAC 78 under alternative solvents and buffers	58
2.4	Stability results of photocaged dAC analogues and dC reference compounds	59
3.1	Summary of locus-specific methylation analysis techniques	77
4.1	Summary of half-lives from MS stability studies	122
A.1	Quantum yield calculations of photocaged analogues	155
A.2	Plate setup for dAC and photocaged dAC cell treatments	183
A.3	Preparation of treatment dilutions for cell experiments	183
A.4	Normalised DNMT1 band intensities in 3'-CD treated Saos-2 cells . . .	188
A.5	Normalised DNMT1 band intensities in 5'-CD treated Saos-2 cells . . .	189
A.6	Normalised DNMT1 band intensities in 3'-CD treated T24 cells	189
A.7	Normalised DNMT1 band intensities in 5'-CD treated T24 cells	190
A.8	Cell seeding densities for low concentration treatments	192
A.9	Preparation of dAC and photocaged dAC dilutions for low concentration treatments	192
A.10	Cell viability treatment setup I.	193
A.11	Preparation of dAC and photocaged dAC dilutions for for high concentration treatments	193
A.12	Cell viability treatment setup II.	194
A.13	Preparation of positive control dilutions for ELISA assay	194
A.14	Assay setup for ELISA	195
A.15	Absorbance readings from first ELISA trial	195
A.16	Absorbance readings from second ELISA trial	196

A.17 LC-MS method conditions	197
A.18 MRM optimisation results	198
A.19 Summary of LC-MS/MS measurements of stability studies	201

List of Schemes

1.1	Structures of non-nucleoside DNMT inhibitors	13
1.2	Structures of first generation azanucleosides	14
1.3	Structures of dAC anomers	19
1.4	Hydrolytic instability of dAC in aqueous solution	21
1.5	Structures of second generation nucleoside DNMT inhibitors	22
1.6	Structures of azanucleoside prodrugs	23
1.7	Principles of photocaging	26
1.8	Popular NB and NPE phototags in literature	28
1.9	Photodeprotection mechanism of NPE-derived carbamates	29
1.10	Structures of coumarin phototags	30
1.11	Synthetic strategies towards coumarin-caged compounds	32
1.12	Photodeprotection mechanism of coumarin-caged molecules	33
1.13	Nucleoside photocaging strategies at 5'- position I.	35
1.14	Nucleoside photocaging strategies at 5'-position II.	35
1.15	Nucleoside photocaging strategies at 3'-position	37
1.16	Nucleoside photocaging strategies at the nucleobase	38
2.1	Synthesis of NPE-chloroformate 58	42
2.2	Synthesis of bis-NPEOC-AC 59	43
2.3	Synthesis of mono-NPEOC-dAC 60	43
2.4	Synthesis of DMNPE-chloroformate 63	44
2.5	Synthesis of mono-DMNPEOC-dAC 65	44
2.6	Uncaging mechanism of bis-NPEOC-AC 59	46
2.7	Synthesis of BHC phototag 71	49
2.8	Hydroxyl-protection of BHC phototag 71	50
2.9	Synthesis of DEACM-chloroformate 77	51
2.10	Synthesis of N-DEACMOC-dAC 78	52
2.11	Proposed mechanism for dark instability of N-DEACMOC-dAC 78	55

2.12	Structures of photocaged dC reference compounds 80 and 81	58
2.13	Proposed mechanism of instability in N-DEACMOC-dAC 78	59
2.14	Synthesis of 5'-DEACMOC-Thymidine 83	61
2.15	Direct coupling of DEACMOC-Cl to dAC	61
2.16	Synthesis of 5'-caged dC <i>via</i> protection of exocyclic amine	63
2.17	Synthesis of DEACMOC-PFP 90	64
2.18	Synthesis of 5'-DEACMOC-dAC 84 <i>via</i> DMAP coupling and PFP activation	65
2.19	Synthesis of 3'-DEACMOC-dAC 93	68
2.20	Proposed stabilising interactions within 5'-DEACMOC-dAC 84	71
2.21	Structures of novel photocaged azanucleoside analogues	74
4.1	Reported stabilities of carbonate linked nucleosides in literature	112
4.2	Possible mechanisms of instabilities in photocaged dAC	113
4.3	Structures of compounds used for stability studies	114
4.4	Mechanism of carbonate link hydrolysis	120
4.5	Proposed mechanism of degradation of photocaged dAC	124
5.1	Structures of non-hydrolysable Azanucleosides	128
5.2	Synthesis of Trityl-dAC 96	129
5.3	Synthesis of DMT-dAC 97	129
6.1	Carbonate/carbamate containing nucleoside and non-nucleoside analogues	139
6.2	Alternative caging strategy via carbamate linkers	142
6.3	Self immolative linker strategy	143
6.4	Suggested analogues with improved hydrolytic stabilities	148

Acronyms

4-NPEOC	4-nitrophenethyl carboxy
5caC	5-carboxylecytosine
5fC	5-formylcytosine
5mC	5-methylcytosine
AC	5'-azacytidine
ACN	acetonitrile
ALU	<i>Arthrobacter luteus</i>
AML	acute myeloid leukaemia
ATP	adenosine triphosphate
BHC	6-bromo-7-hydroxymethylcoumarin
BPFPC	bis(pentafluorophenyl)carbonate
CBz	carboxybenzyl
CDA	cytidine deaminase
CES1	Carboxylesterase 1 enzyme
CML	Chronic Myeloid Leukaemia
CpG	Cytosine-phosphate-Guanine
dAC	5-aza-2'-deoxycytidine
DBU	1,8-diazabicycloundec-7-ene
dCK	deoxycytidine kinase
dCMPK	deoxycytidine monophosphate kinase
DEACM	7-diethylaminomethylcoumarin
DIAD	diisopropyl azodicarboxylate
DIPEA	<i>N,N</i> -diisopropylethylamine
DMAP	4-(dimethylamino)pyridine
DMNPE	4,5-dimethoxy-2-nitrophenethyl
DNMT	DNA methyltransferase
DPK	diphosphokinase

GAPDH	glyceraldehyde 3-phosphate dehydrogenase
hCNT	human concentrative nucleoside transporter
hENT	human equilibrative nucleoside transporter
HMDS	hexamethyldisilazane
HPCE	high performance capillary electrophoresis
HPLC	high performance liquid chromatography
LINE-1	long interspersed nuclear elements 1
LSD1	Lysine Specific Demethylase 1
LUMA	luminometric methylation assay
MAGE	melanoma antigen
MDS	myeloid dysplastic syndromes
MOM	methoxymethyl
MOM-Cl	methoxymethyl chloride
NB	<i>o</i> -nitrobenzyl-
NBMPr	<i>S</i> -(4-nitrobenzyl)-6-thioinosine
NBS	<i>N</i> -bromosuccinimide
PBS	phosphate-buffered saline
PFP	pentafluorophenyl
RRM2	ribonucleotide reductase
SAH	S-adenosyl-homocysteine
SAM	S-adenosyl-L-methionine
SEM	2-(trimethylsilyl)ethoxymethyl
SEM-Cl	2-(trimethylsilyl)ethoxymethyl chloride
TALE	Transcription Activator-Like Effector
TBAF	tetrabutylammonium fluoride
TBDMS	<i>tert</i> -butyldimethylsilyl
TDG	thymine-DNA glycosylase
TET	ten-eleven translocation
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TIPS	triisopropylsilyl
TIPS-Cl	triisopropylsilyl chloride
TLC	thin layer chromatography
ZFN	Zinc Finger

Chapter 1

Introduction

1.1 Epigenetics

DNA is one of the most important biological molecules as it carries genetic information. Nucleotides are the smallest building blocks of DNA, containing a phosphate linker, deoxyribose sugar and one of four nitrogen-containing aromatic bases: guanine (G), adenosine (A), thymine (T) and cytosine (C). The sequence of nucleotides encodes the genetic information in coding regions of genes, which are transcribed into triplet codons and further translated into corresponding amino acid sequences.¹ The latter carry vital instructions on how proteins are synthesised and function in living organisms.² The haploid human genome contains around 3 billion nucleotides and around 20 000 - 25 000 protein-coding genes, but these genes are utilized differently across various cells and developmental stages. The genetic code is essentially a script of all the information a cell *could* use to function, but its use depends on how the cell reads this script.³

Epigenetics determines how the genetic code is read by the cell. The word epigenetics is of Greek origin, meaning 'above' or 'on top' of genetics.⁴ It refers to the mitotically heritable changes on top of the DNA that do not alter the nucleotide sequence, but highly influence gene expression.⁵

A classic example in which epigenetics plays an essential role is in developmental biology.⁶ Waddington's illustration of the epigenetic landscape (Figure 1.1) serves as a metaphor of how gene regulation modulates development.⁷ In a multicellular organism, all cells are 'born' with the same genetic material. The marble on the top of the hill represents a pluripotent cell that has the potential to differentiate into any cell type.⁸

During cell development, different sets of genes are switched on as a result of epigenetic reprogramming and this triggers differentiation. As the cell moves towards its final differentiated state, represented by the movement down the hill along a restricted path, its potential to become a different cell type diminishes.

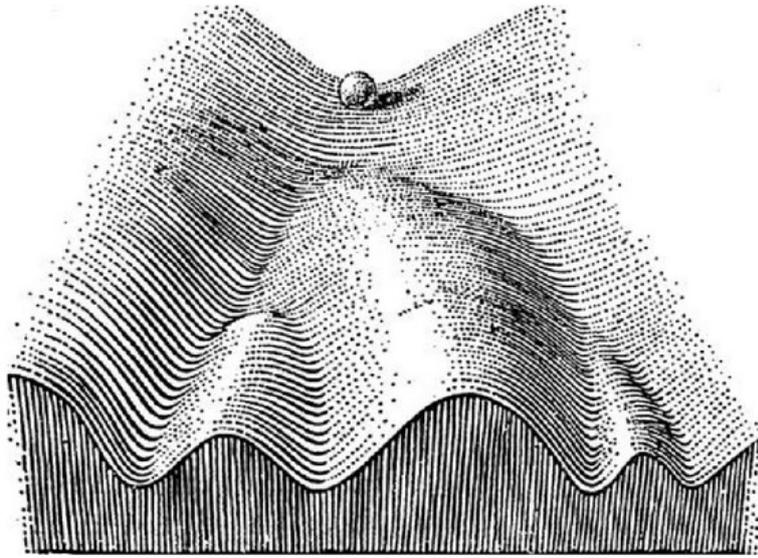


Fig. 1.1 The epigenetic landscape by Conrad Waddington⁸

The expression of a gene is determined by the accessibility of the DNA to the transcriptional machinery.⁹ DNA is packaged and wrapped around an octamer of histone proteins to form a nucleosome unit, which is further organised into higher organisational units of chromatin.¹⁰ Chromatin can either exist in relatively open, accessible and transcriptionally active domains (euchromatin) or more compact and inactive forms (heterochromatin).¹¹ Epigenetic modifications maintain a balance between these two states by either remodeling or directly altering the chromatin structure in a covalent or non-covalent fashion. The three main modifications include: histone modifications, non-coding RNAs and DNA Methylation (Figure 1.2).¹²

The post-translational modifications of histone proteins (such as acetylation, methylation, phosphorylation, ubiquitination and sumoylation) can greatly influence the biophysical properties of histones, which have downstream effects on the chromatin structure.¹³ For example acetylation of lysine residues removes the positive charge on histones, which decreases interaction with the negatively charged DNA. This loosens the structure of the condensed chromatin and as a result, DNA becomes more accessible to transcription. Nevertheless, following successful transcription, gene expression can

still be blocked. MicroRNAs selectively bind to target messenger RNAs (mRNAs) with a complementary sequence to induce cleavage or degradation and subsequently block translation and gene expression.¹⁴ These epigenetic modifications have been well reviewed in literature. However, DNA methylation is probably the most extensively studied epigenetic mark.^{15,16}

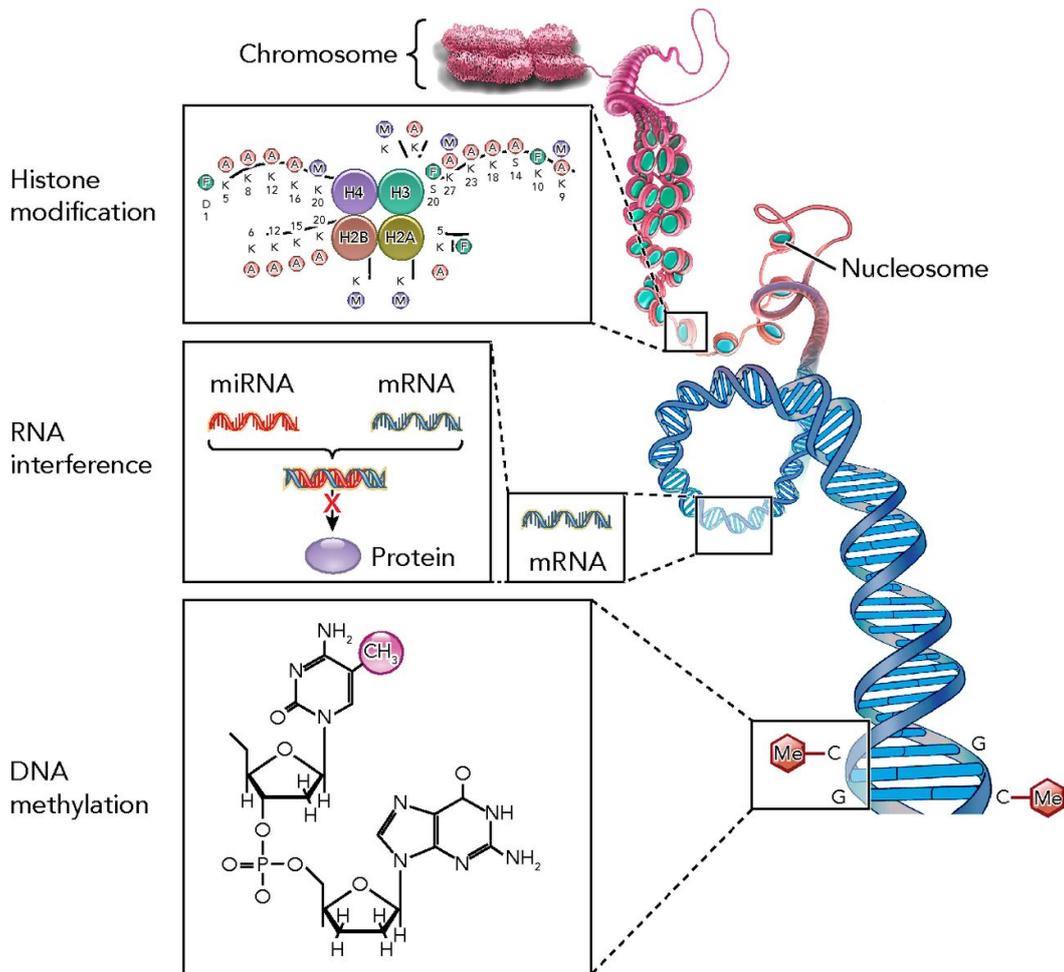


Fig. 1.2 Structure of the chromosome and summary of epigenetic modifications¹⁷

1.2 DNA methylation

1.2.1 DNA methylation in mammals

In mammalian DNA, methylation occurs predominantly on Cytosine-phosphate-Guanine (CpG) dinucleotides, where the cytosine residue is followed by a guanine nucleotide.¹⁸ If located within a promoter sequence, the presence of a methyl group is associated with transcriptional silencing.¹⁹ The methyl group can physically block the binding of transcription factors. In addition, it can also recruit methyl-CpG-binding proteins (MBPs), which can further recruit transcriptional co-repressor molecules to induce downstream gene repression.²⁰

Within the genome, CpG dinucleotides are not distributed evenly. They are concentrated in short CG rich stretches of DNA ('CpG islands') and regions of long repetitive sequences such as centromeric repeats.²¹ CpG islands are found in proximal promoter regions of at least 60% of mammalian genes. Within these regions, CpGs are unmethylated in order to be recognised by transcription factors and associated DNA-binding proteins. Methylated CpGs within promoter regions can also occur naturally, for example during X-chromosome inactivation²² and genomic imprinting.²³ On the contrary, CpGs found within the rest of the genome tend to be heavily methylated in order to prevent chromosomal instability. Methylation in these regions prevents the expression of unwanted repetitive sequences and transposable DNA elements.²⁴

1.2.2 DNA methyltransferases

DNA methylation patterns are established and maintained by the family of DNA methyltransferase (DNMT) enzymes.²⁵ While five different DNMT isoforms exist, only three are active in mammalian cells: DNMT1, DNMT3A and DNMT3B. DNMT3A and DNMT3B are essential for the establishment of new methylation patterns (*de novo* methylation).²⁶ Targeted disruption of these enzymes resulted in lethality in mice.²⁷

Maintenance of global methylation patterns is achieved by DNMT1 enzyme and its binding partner Ubiquitin-like Homeodomain and RING Finger domain 1 (UHRF1).^{28,29} These enzymes are ubiquitously expressed in proliferating cells. They are localized in the replication foci during S-phase and preferentially bind to hemimethylated CpGs, where only one of the two strands in the parent duplex is methylated.³⁰ For each hemimethylated CpG site, the target recognition domain of DNMT1 first recognizes

and binds to the parent strand and the catalytic domain then binds to the corresponding cytosine on the growing daughter strand.²⁸ This is followed by the covalent transfer of a methyl group onto the cytosine of the newly replicating strand.

The catalytic mechanism of DNMT is illustrated in Figure 1.3. The DNMT enzymes catalyse the covalent transfer of a methyl group from S-adenosyl-L-methionine (SAM) cofactor molecule onto position 5 of the cytosine base (a to e).³¹ This transfer consists of several steps. Firstly, a covalent complex is formed between the DNMT1 catalytic pocket and the cytosine residue (b). This is initiated by the nucleophilic attack of the highly conserved cysteine C1229 on position 6 of the cytosine base. The adjacent glutamic acid residue (Glu1266) can simultaneously stabilize the increased electron density of the aromatic ring by proton donation to the nitrogen atom at position 3 (c). In the presence of the positively charged SAM cofactor, the electron-rich cytosine base attacks the methyl group, which is subsequently released from the cofactor (c). As a result, neutral S-adenosyl-homocysteine (SAH) is formed and the methyl group becomes covalently bound to the C5 of the cytosine base. The final step of the transfer is the β -elimination of the proton and cysteine residue at position 5 and 6 respectively (d). This step enables the regeneration of the cytosine aromatic system and the release of the DNMT1 enzyme along with the 5-methylcytosine (5mC) product.

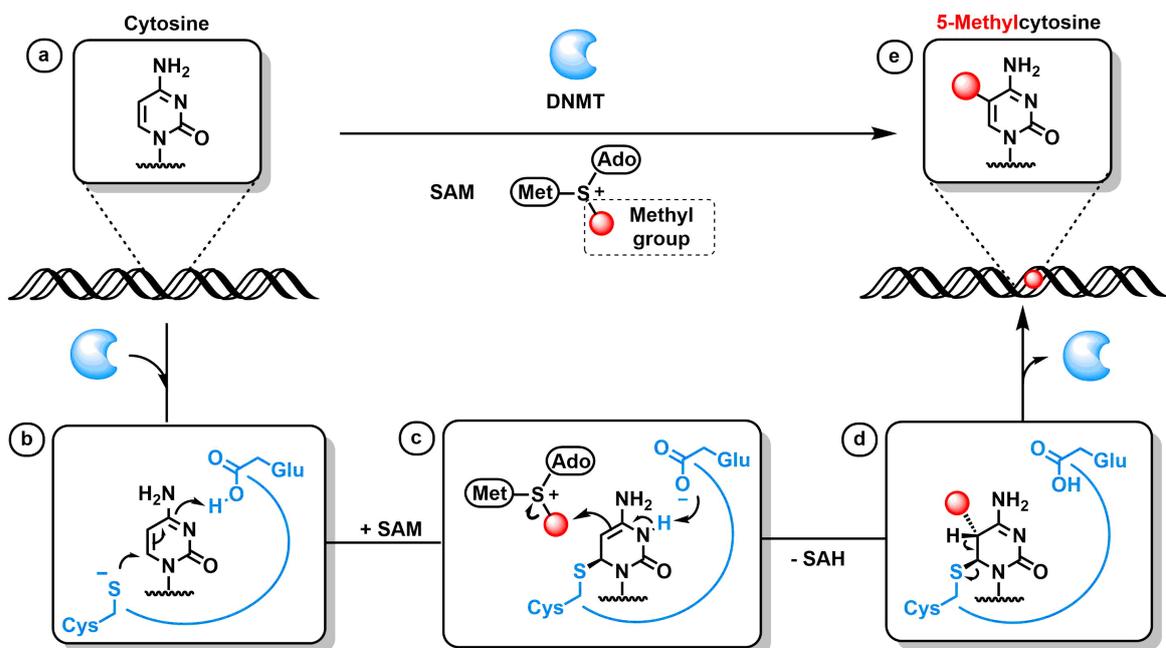


Fig. 1.3 Catalytic mechanism of cytosine methylation by DNMT1

While the mechanism of DNA methylation is well understood, considerably less was known about DNA demethylation. Passive DNA demethylation occurs, when maintenance methyltransferases are inactivated during the replication cycle.³² As a result, the unmethylated state is retained in the newly synthesised daughter strands. This process depends on DNA replication and thus takes time to occur. Conversely, active demethylation is more rapid; it involves one or more ‘demethylase’ enzymes that operate independently of DNA replication.

1.2.3 TET enzymes

The recent discovery of the ten-eleven translocation (TET) enzymes has provided major insight into the mechanism of active DNA demethylation.^{33,34} It has been found that TET enzymes can oxidise 5mC to 5hmC in the presence of Fe(II) and α -ketoglutarate (α -KG).³⁵ The proposed catalytic pathway is shown in Figure 1.4. At the active site of TET, Fe(II) is bound by highly conserved His-His-Asp residues, water and α -KG. In the presence of oxygen, the α -KG is oxidised into enzyme-bound succinate and the Fe(II) complex results in a high-valent Fe(IV)-oxo intermediate. The intermediate reacts with 5mC to yield 5hmC *via* a net oxidative transfer of single oxygen. As a result, the Fe(II) species is regenerated at the end of the reaction.³⁵

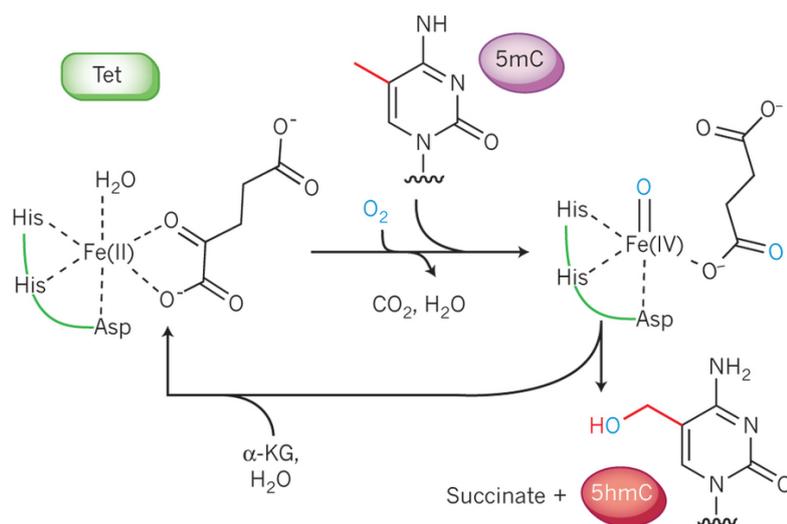


Fig. 1.4 Catalytic mechanism of Fe and α -KG dependent oxidation of 5mC to 5hmC by TET³⁵

The 5hmC can be further oxidised to 5-formylcytosine (5fC), 5-carboxylcytosine (5caC) which are then excised by thymine-DNA glycosylase (TDG) enzymes to generate an abasic site.³⁶ These sites are targeted by the base excision repair process to regenerate cytosine at the site of interest.³⁷ The dynamic modification cycle of cytosine analogues by DNMT, TET and TDG enzymes is summarised in Figure 1.5.³⁵ The molecular details of these conversion processes and the regulatory roles of these analogues are currently being investigated.^{38–41}

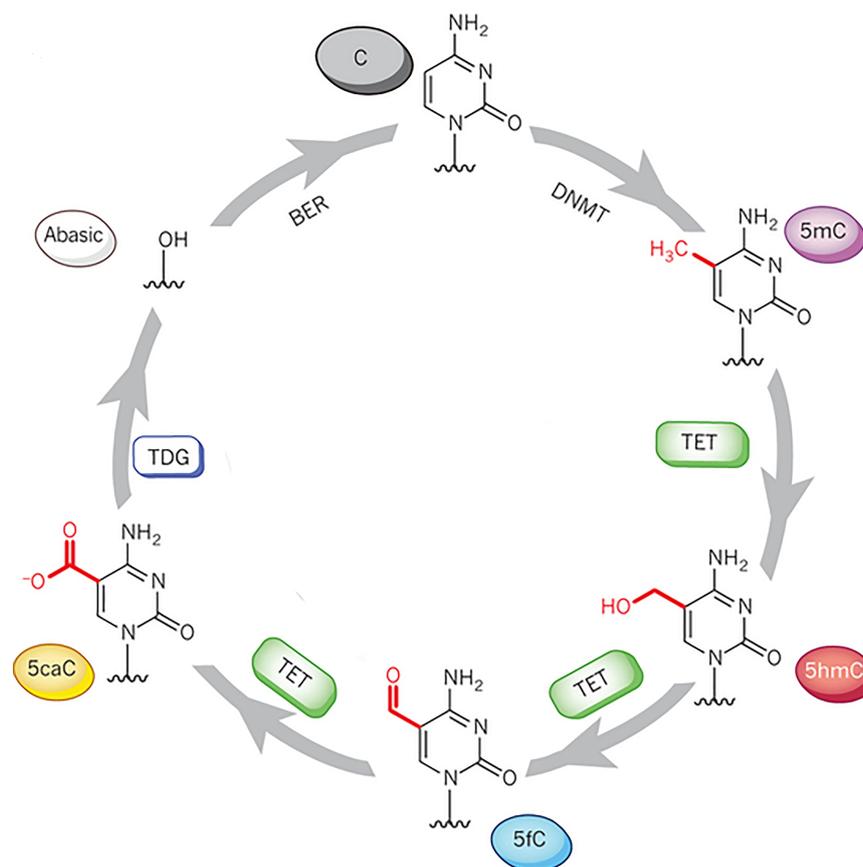


Fig. 1.5 Dynamic modifications of cytosine analogues³⁵

Three different TET isoforms have been discovered to date: TET1, TET2 and TET3. Despite their similar catalytic activity, each isoform has a distinctive function in development. TET1 plays an important role in maintaining pluripotency in embryonic stem (ES) cells.⁴² It has been shown that active demethylation is important for resetting the epigenetic state of the parental genome.⁴³ For example in early embryos, parental-origin-specific imprints are erased in order to set up pluripotent states.⁴⁴ TET2 on the other hand functions as a physiological regulator of haematopoiesis,⁴⁵ whereas TET3 acts as a critical factor for zygotic epigenetic reprogramming.⁴⁶

1.2.4 Disruption of methylation pattern in disease

DNA methylation and demethylation play a crucial role in cell development,⁴⁷ regulation of key cellular processes⁴⁸ and maintaining chromosome stability.^{49,50} Consistent with these vital roles, perturbations of DNA methylation patterns are frequently observed in a growing number of diseases. The diseases affected by aberrant DNA methylation include imprinting disorders, chromosomal instabilities and cancer.⁵¹

Imprinting disorders arise when genomic imprinting is lost during development. Genomic imprinting is a phenomenon whereby certain genes are expressed in a parent-of-origin-specific manner.²³ This is due to epigenetic modifications that lead to silencing of a specific parental allele in the gamete or zygote. As a result, only the other parental allele is expressed in the somatic cells of the offspring. Although genomic imprinting only affects around 100 genes in mammals, there are a number of human diseases that are associated with disruptions of imprinted epigenetic marks through gain or loss of DNA methylation. Examples include Angelman's Syndrome, Prader-Willi syndrome⁵² and Beckwith-Wiedemann syndrome.⁵³ In these conditions an abnormal phenotype is established as a result of missing parental copies or the deregulation of imprinted genes.⁵⁴

Chromosomal instabilities can emerge from global changes in DNA methylation levels.⁵⁵ Such alterations are the direct results of mutational changes in the genomic regions of regulatory enzymes. For example, ICF syndrome (Immunodeficiency, Centromeric region instability and Facial anomalies) is caused by mutations in the DNMT3B gene that are linked to abnormal global DNA hypomethylation.^{56,57} Furthermore, Rett Syndrome - a severe neurological disorder characterised by developmental regression and seizures - is a result of germline mutations of MeCP2 methyl binding protein.⁵⁸ The role of MeCP2 protein is to bind to methylated DNA and initiate gene silencing. In Rett Syndrome, the protein is unable to bind to DNA and silence genes. As a result, genes that are normally repressed by MeCP2 remain active when their products are not needed. This ultimately leads to the disruption of the normal functioning of nerve cells.⁵⁸

Cancer is one of the most significant diseases in which the disruption of DNA methylation patterns can have a crucial role. The link between aberrant DNA methylation and cancer has been established several decades ago,⁵⁹⁻⁶¹ but the underlying cause of methylation defects in cancer is still unknown. It is now widely recognised that aberrant DNA methylation is perhaps the most common molecular lesion found in

cancer cells.⁵⁵ These lesions are generally the result of two aberrant methylation events: hypomethylation within the genome and hypermethylation at gene promoter regions as shown in Figure 1.6.⁶²

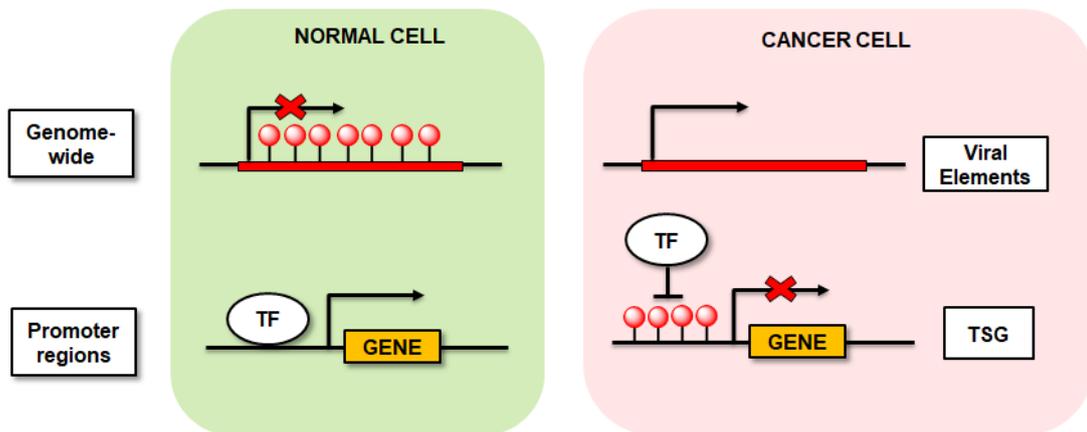


Fig. 1.6 DNA methylation patterns in normal and cancer cells⁶²

Loss of DNA methylation at CpG dinucleotides was the first epigenetic abnormality identified in cancer cells.^{63,64} These events occur within the genome and result in illegitimate recombination events and expression of potentially damaging and viral elements.⁶⁵ While loss of DNA methylation generally has global consequences *via* predisposing cells to genetic instability, gene-specific effects of hypomethylation may also occur. For example, the melanoma antigen (MAGE) family of genes are normally silenced in healthy cells. They encode tumour antigens and are frequently demethylated and re-expressed in cancer.

Hypermethylation of gene-promoter regions is the other aberrant methylation event that is frequently observed in cancer. The resulting changes effectively silence the expression of key tumour suppressor genes (TSG) involved in cell-cycle regulation, cell signalling, DNA repair and tumour suppression.^{66,67} Examples of frequently hypermethylated genes are summarised in Table 1.1.⁶⁸ In addition, apoptotic genes are also known to become silenced in nearly every tumour type.⁶⁹ Such changes provide tumour cells great advantage in survival, facilitating uncontrolled growth and proliferation.

Gene	Function	Tumour type
Rb	Cell-cycle regulation	Retinoblastoma
APC	Wnt signal transduction	Colorectal cancer
p14/ARF	Cell-cycle regulation	Colorectal cancer
p15/CKDN2B	Cell-cycle regulation	Leukemias
p16/CDKN2A	Cell-cycle regulation	Various cancers
BRCA1	DNA repair	Breast, ovarian cancer
VHL	Tumour suppressor	Renal cell cancer
hMLH1	DNA mismatch repair	Gastric, endometrial cancers
ER- α	Estrogen receptor- α	Breast, colorectal cancers

Table 1.1 Frequently hypermethylated genes in sporadic cancers

1.3 Strategies to target DNA methylation

Given that aberrant DNA methylation is associated with a great number of diseases, the regulatory enzymes of this machinery such as DNMTs are seen as attractive therapeutic targets.⁷⁰ In contrast to most genetic mutations, epigenetic alterations are reversible and it is possible to reprogram them.⁷¹ Reversing methylation is not only of therapeutic interest, but it could also serve as a research tool. The ability to experimentally lower methylation levels could provide useful insights on the cellular processes that are normally regulated via DNA methylation patterns.

Currently there are two main strategies to reverse methylation levels in biological organisms. Methylation can be targeted in a locus-specific manner using biological editing tools or it can be targeted in a nonsequence-specific (global) manner using small molecule chemical tools. For the latter strategy, only lowering of methylation levels have been achieved to date, whereas biological tools have been successfully used to add or remove methylation marks.

1.3.1 Biological editing tools

Several biological tools have been developed to date that can achieve locus-specific epigenome editing (Summarised in Figure 1.7).^{72,73} Together these tools enable the selective activation or repression of specific genes. These techniques in combination with comparative phenotypic analyses can further aid our understanding of how individual epigenetic modifications can exert control on processes in the cellular machinery.

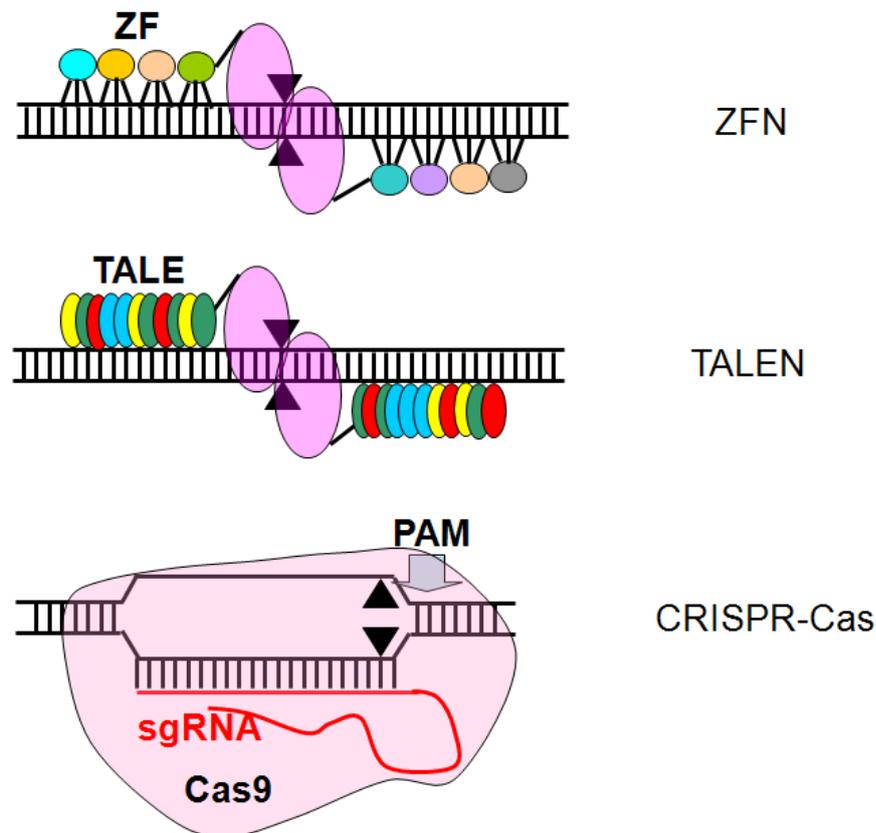


Fig. 1.7 Models of biological editing tools⁷⁴

The classical epigenome editing technology relies on custom-engineered proteins that are composed of two modules: a DNA binding domain, which can target specific sequences of the genome, and an effector domain that modifies epigenomic features (Figure 1.7). The Transcription Activator-Like Effector (TALE) and Zinc Finger (ZFN) proteins are often chosen as DNA binding domains due to the relative ease to engineer them to target different genomic regions of interest.^{75,76} DNA binding domains have been fused to a variety of chromatin modifiers to carry out epigenome editing (Figure 1.7, pink circles). Examples include the TALE-TET1 fusion system that successfully removed

methylation at target sites although with moderate efficiency.⁷⁷ Another example was TALE fused to Lysine Specific Demethylase 1 (LSD1) that removed methylation from histone proteins (H3K4me1 and H3K4me2).⁷⁸ Furthermore, it was also demonstrated that the effector proteins can be recruited with a light-sensitive cryptochrome protein, which introduced an additional spatiotemporal control to the approach.⁷⁹

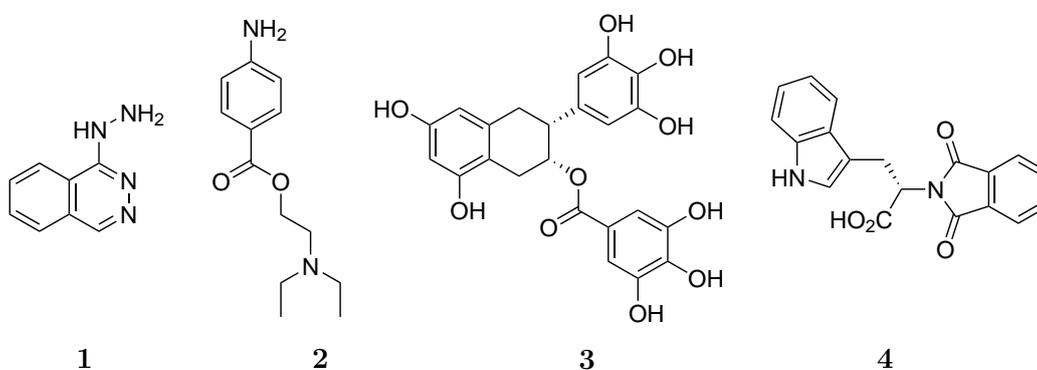
Despite the promising results, the design of TALEs and ZFNs is still cumbersome and labour intensive. The recent development of the CRISPR/Cas9 genome editing technology initiated a new wave of excitement within the scientific community because it offered a faster, cheaper and more accurate approach to edit DNA.^{80,81} CRISPR stands for Clustered Regularly Interspaced Short Palindromic Repeats. The CRISPR/Cas9 system is composed of 3 distinct modules: A guide RNA, a modified bacterial Cas9 nuclease and an effector chromatin modifier of choice. The guide RNA sequence determines the binding position of the complex within the genome and it can be custom-designed at a low cost.⁸² The catalytically dead dCas9 nuclease is fused to a chromatin modifier and the complex together with the guide RNA are used to add or remove epigenetic marks, depending on the chromatin modifier. Several groups have successfully fused the CRISPR/Cas9 system with DNMT3A or TET1 enzymes to add or remove DNA methylation marks at specific loci in the genome with improved efficiencies compared to TALE systems.⁸³⁻⁸⁵

The TALE and CRISPR/Cas9 systems enable the interrogation of individual sites in the genome with precision at the nucleotide level. While these tools are invaluable for our understanding of epigenetic modifications, the field is still in its infancy - there is still room for optimisation of modification efficiencies. In spite of this, the results of the first trials on human embryonic cells have recently been reported by Ma and coworkers.⁸⁶

While biological tools enable the targeting of individual methylation marks, small molecules provide a genome-wide approach to induce methylation changes. Small molecules are easier to prepare and modify compared to biological tools. To date, several agents have been developed in order to reverse DNA methylation through direct or indirect interference with DNMT enzymes.⁸⁷ Inhibition of DNMTs result in the blocking of maintenance methylation of newly replicating DNA strands, leading to genome-wide lowering of methylation levels in daughter strands (global demethylation). A number of analogues have high 'demethylating' efficacy and long-established clinical profiles. Chemical tools can be classified into non-nucleoside inhibitors and nucleoside inhibitors.

1.3.2 Chemical tools I: non-nucleoside inhibitors

Non-nucleoside DNMT inhibitors can directly interfere with the enzyme activity. Although in most cases the mechanism of inhibition is unknown, a number of small molecules have been reported to date that showed moderate inhibitory effect against DNMT1 (Scheme 1.1). Among drugs known for different applications, hydralazine **1** and procainamide **2** – known cardiovascular drugs – have been reported to inhibit DNA demethylation in cells.⁸⁸ (-)-Epigallocatechin3-gallate **3** a major component of green tea extracts has also been shown to inhibit DNA methylation presumably through binding to DNMT1.⁸⁹



Scheme 1.1 Structures of non-nucleoside DNMT inhibitors

RG108 **4** is a small-molecule inhibitor of DNMT1 that was developed using a computational screening approach based on the three-dimensional structure of DNMT1.⁹⁰ This analogue has shown promising activity *in vitro* by inhibiting methylation of tumour suppressor genes at gene promoter regions. In addition, it does not demonstrate cytotoxic effects even at high concentrations.⁹¹ Various derivatives of RG108 are currently being investigated to identify more potent inhibitors.⁹²

The final and possibly most promising non-nucleoside DNMT inhibitor is MG98. It is a second-generation 20mer antisense oligonucleotide designed to hybridize with DNMT1 mRNA and prevent further processing of the mRNA and eventually reduce cellular levels of DNMT1.⁹³ There are two particular chemical traits that enhance the plasma life of MG98. First of all, the conventional phosphodiester linkages are replaced by phosphorothioate linkages which improve its stability in plasma. Secondly, specific 2'-*O*-methyl modifications are included on the strands in order to increase the molecule's affinity for mRNA targets and reduce off-target binding and toxicity. MG98

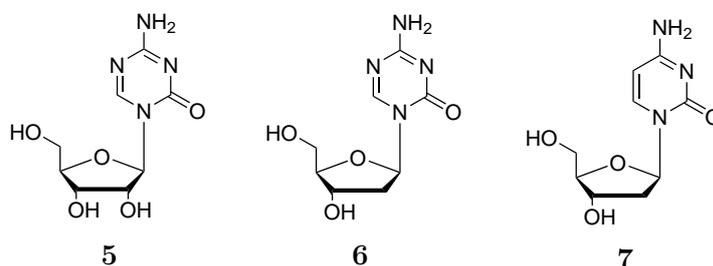
has been extensively investigated clinically in advanced solid tumours and renal cell carcinomas.⁹⁴

Non-nucleoside inhibitors offer a strategy to selectively target DNMT1 enzymes, however the potency of these analogues need further improvement. Conversely, nucleoside inhibitors are historically the most popular strategies to target and reverse DNA methylation in organisms.

1.3.3 Chemical tools IIa: nucleoside inhibitors

Introduction: Azanucleosides

The two most widely used nucleoside DNMT inhibitors to date are the azanucleosides 5'-azacytidine (AC) (VidazaTM, Azacitidine) **5** and 5-aza-2'-deoxycytidine (dAC) (DacogenTM, Decitabine) **6** (Scheme 1.2).⁹⁵ They are structural analogues of 2'-deoxycytidine **7** that contain a nitrogen in place of the 5-carbon of the pyrimidine ring.



Scheme 1.2 Structures of first generation azanucleosides and cytidine

The history of azanucleosides dates back to the 1960s. They were first synthesised by Pískala and Šorm.⁹⁶ Originally intended to be used as cytotoxic agents, they were tested in leukaemia therapies,^{97,98} until their demethylating effect was reported in 1980.^{99,100} AC and dAC were clinically approved in 2004 and 2006 respectively for the treatment of myeloid dysplastic syndromes (MDS).^{101,102} In addition, AC is also used in the treatment of acute myeloid leukaemia (AML).

The molecular pharmacology of azanucleosides

The mechanisms of uptake and activation of azanucleosides are shown in Figure 1.8.¹⁰³ Azanucleosides enter the cells via nucleoside transporters that normally allow the uptake of cytidine and uridine.¹⁰⁴ Nucleoside transporters are classified into two main families, the Na⁺-dependent human concentrative nucleoside transporter (hCNT) and the Na⁺-independent human equilibrative nucleoside transporter (hENT) proteins. Among the hCNT family isoforms, isoform 1 (hCNT1) has been shown to have the highest affinity to AC ($K_m = 63 \mu\text{M}$), but isoform 3 (hCNT3) showed the highest rate of transport.^{104,105} Other transporters in the hENT family also showed modest transport capacities for AC.

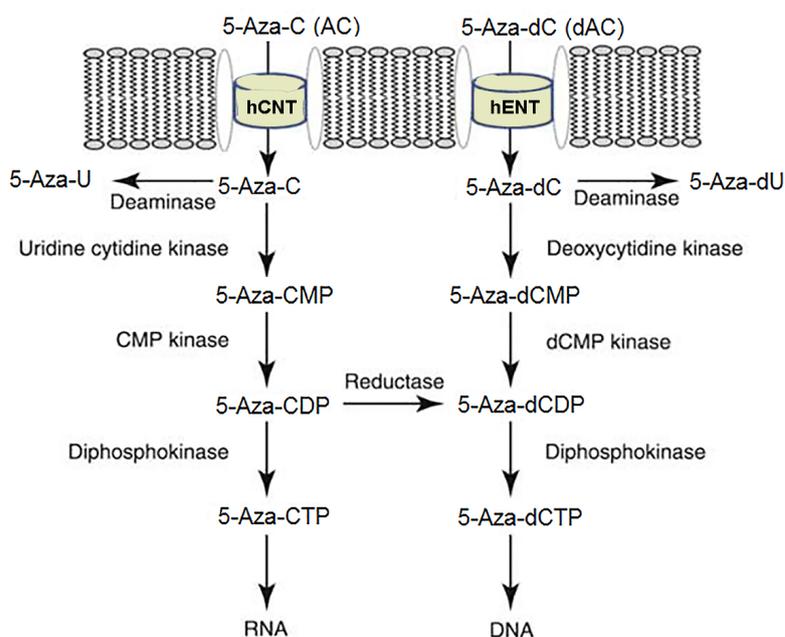


Fig. 1.8 Uptake and incorporation of azanucleosides¹⁰³

Conversely, dAC is mainly transported by hENTs (isoform 1 and 2). The transport of dAC - similar to uridine - is concentration and time-dependent and Na⁺-independent.¹⁰⁴ This suggests facilitated-diffusion uptake and is a characteristic of hENT transport. The K_m value for dAC uptake by hENT1 is $1.09 \mu\text{M}$, which is 900 times lower than that of uridine to hENTs and 60 times lower than the affinity of AC to hCNT1.¹⁰⁶ The transport rate of dAC by hENTs is also substantially lower compared to the transport of AC by hCNTs.

In both cases, AC and dAC transport can be mitigated by the presence of the nucleoside transport inhibitor *S*-(4-nitrobenzyl)-6-thioinosine (NBMPR).^{104,106} In summary, the presence, absence and types of nucleoside transporters (for example hCNT1 in kidney cells or hENT1 in colon carcinoma (HCT116) cells) are key determinants of the activity of azanucleosides in the cellular environment.

Once inside the cell, azanucleosides are converted into their active form to become substrates for the DNA replication machinery. This is achieved *via* three consecutive steps of phosphorylation by kinases (Figure 1.8). In the case of dAC, the first step is catalysed by deoxycytidine kinase (dCK) to produce the monophosphorylated form of dAC, or 5-aza-2'-deoxycytidine-5'-monophosphate (5-aza-dCMP).¹⁰⁷ This is the rate-limiting step of the activation. The two subsequent steps are catalysed by deoxycytidine monophosphate kinase (dCMPK) and the nucleoside diphosphokinase (DPK) to produce the di- and triphosphorylated version of the drug respectively.¹⁰⁸ The activated drug then gets incorporated into replicating DNA with the aid of DNA polymerases.¹⁰⁹

In a similar fashion, AC is phosphorylated by three different kinases, and in addition 10% of it is converted into dAC analogues *via* a reductase enzyme.¹¹⁰ However, the majority of activated AC (80-90%) is incorporated into RNA in order to interfere with protein synthesis.¹¹¹ Azanucleosides belong to the family of S-phase specific drugs, suggesting that they are only active once incorporated in the DNA/RNA of cells.¹¹² This enables them to selectively get incorporated into the DNA/RNA of rapidly dividing cells, which gives them a preference towards cancer cells.¹¹³

Azanucleosides, once incorporated into the DNA, are recognised by DNMTs, similarly to cytidines. However, the 5-nitrogen has a key role in the mechanism of inhibition as illustrated in Figure 1.9.

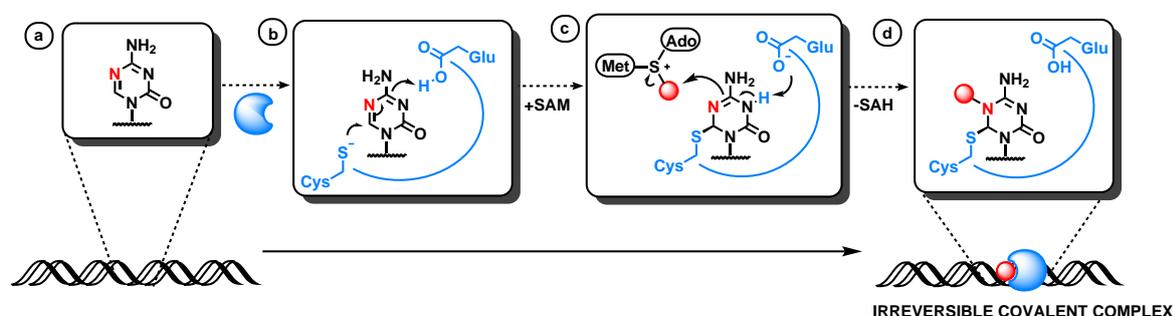


Fig. 1.9 Mechanism of DNMT1 inhibition by azanucleosides

The nitrogen essentially blocks the final β -elimination step due to the absence of a proton in position 5. As a result, the enzyme remains covalently bound to the DNA and its demethylating function is blocked. The resulting covalent complex not only prevents DNMTs from methylating further DNA strands in replicating DNA, but it is also targeted by ubiquitin-E3 ligase for proteosomal degradation.¹¹⁴ This leads to depletion of DNMT1 levels and the subsequent reduction in DNA methylation levels (global genome-wide demethylation) in daughter strands.¹¹¹ This is the commonly accepted mechanism of azanucleosides, which is based on the initial report by Santi and coworkers.¹¹⁵ It states that the activity of dAC is replication dependent, suggesting that dAC is required to be incorporated in order to exert DNMT1 depletion.¹¹⁶

However, a study by Ghoshal and coworkers suggested that there may be an additional pathway towards DNMT depletion by dAC that is replication-independent.¹¹⁷ Their study suggests that dAC is able to inhibit free DNMT1 through interaction with Protein kinase C δ (PKC δ) which then hyperphosphorylates DNMT and induces proteosomal degradation.¹¹⁸ The exact mechanism of this phenomenon remains elusive.

Effect of azanucleosides on RNA metabolism

Numerous studies have indicated that the mode of action of azanucleosides are not limited to DNA methylation inhibition. It has been shown that incorporation of triphosphorylated azacytidine into RNA inhibits tRNA methylation and processing by reducing tRNA methyltransferase levels.¹¹⁹ This can subsequently give rise to defective messenger and transfer RNA. The disrupted tRNA processing can ultimately lead to inhibition of mRNA and protein synthesis and thus inducing apoptosis.¹²⁰

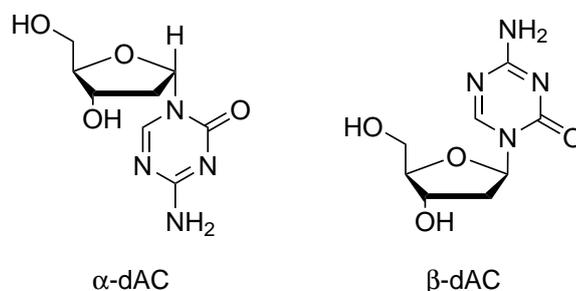
It has also been shown that azacytidine incorporation into RNA is able to repress the expression of the M2 subunit of ribonucleotide reductase (RRM2) and therefore interfere with the conversion of ribonucleotides to deoxyribonucleotides, ultimately leading to inhibition of DNA synthesis and repair.¹²⁰

Mechanisms of resistance to azanucleosides

The activity of azanucleosides greatly depends on each of these systems described above (uptake, phosphorylation, incorporation and inhibition of DNMT1). When one of these components is altered, resistance against dAC can develop.¹²¹ There are three main pathways that are closely correlated with dAC resistance that have been identified and characterised to date:

1. Low hENT expression by the cell leads to reduced uptake of dAC¹²²
2. Low dCK expression or deficiency leaves the cell unable to activate dAC¹²¹
3. Enzymatic degradation by cytidine deaminase (CDA) renders dAC inactive by converting it into 5-azauridine compounds.¹²³ Efforts were made to circumvent this by using CDA inhibitors such as tetrahydrouridine, however the concentration of these compounds need careful consideration as they can cause toxicity at elevated levels.^{124,125}

In addition, the chemical configuration of azanucleosides is essential to their activities. Two structural isomers of dAC have been characterised to date: α - and β -dAC (Scheme 1.3). The isomers differ based on the relative configuration and stereochemistry of the triazine ring to the deoxyribose sugar.



Scheme 1.3 Structures of dAC anomers

The α anomer is more stable compared to the β anomer, however it has a less pronounced bioactivity. This is because the α anomer itself is not incorporated into DNA and is not recognized and degraded by cytidine deaminase.¹²⁶ However, its biological activity is based on its spontaneous conversion into the β isomer which then enters the DNA synthesis pathway.¹²⁷ It has been shown that the α isomer has hypomethylating capacity¹²⁸ and it can also downregulate hTERT mRNA expression in human leukemia cell lines.¹²⁶ Davies *et al.* investigated the crystal structure of dAC and have established that dAC as we know today is mostly composed of the β isomer.¹²⁹

dAC as a demethylating tool

dAC is a powerful demethylating tool.¹³⁰ It is able to reverse hypermethylated states in the genome (regions with high density of methylation) by targeting and inhibiting the methylating enzyme DNMT during replication, which leads to the lowering of methylation levels in the resulting daughter strands. The demethylating function of azanucleosides is most evident at lower drug concentrations (sub 1 μ M) because at higher concentrations these drugs exhibit greater cytotoxicity by interfering with DNA synthesis and causing DNA damage.¹³¹

Ever since the hypomethylating effect of dAC was established, there has been a growing interest in the use of dAC not only as a therapeutic agent, but also as a research tool. The ability to modify methylation levels of cells and analyse the resulting cellular and

phenotypic changes has greatly aided our understanding of the functional significance of DNA methylation in different biological processes.¹³² There are many studies that try to correlate the changes in methylation profiles in response to dAC treatment and resulting changes in gene expression profiles and cellular pathways in different cancer cell lines.^{133–140}

Most of our current knowledge of genetic alterations present in human cancers is derived from *in vitro* established human cancer cell lines.¹⁴¹ The advantages of biological studies in cancer cell lines is that there are a wide spectrum of tumour types available and they are easy to culture *in vitro*. It is relatively easy to get significant amounts of DNA and they are usually free of contamination. However, it is known that different cell lines (even tumour types) have varying baseline methylation levels and gene expression profiles and varying sensitivity to dAC treatment.^{142,143}

To create a normalized database, a systematic profiling study was carried out by Paz and coworkers, whereby the baseline methylation levels of 70 tumour cell lines were measured and their response against dAC treatment was assessed.¹⁴¹ There was considerable variation in percentage CpG island hypermethylation between cell lines (7-73%), with haematological cell lines having the highest and renal/thyroid cells having the lowest methylation. It was also found that dAC treatment left the cells with a narrow range of 5mC content. The more hypermethylated CpG sites a cell line has, the more sensitive it is to dAC treatment.

In addition, dAC was also used for a 'chemogenomic' screening in gastric cancer cell lines.¹⁴⁴ Genes that were silenced in gastric cancer cells were identified and were successfully re-expressed following dAC treatment. Although it is often stated that dAC is a genome-wide hypomethylator, several studies suggested that the functional effects of dAC treatments are limited to a small subset of protein-coding genes and are associated with specific CpG sites in CpG islands.^{145,146}

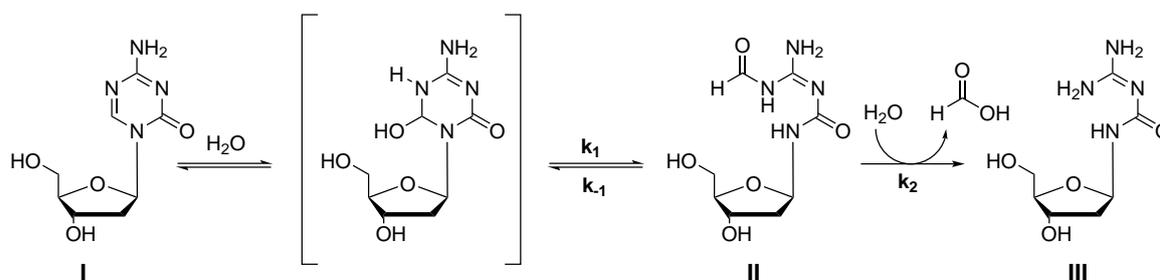
The role of dAC is rapidly emerging in other fields such as regenerative medicine,¹⁴⁷ developmental biology¹⁴⁸ and cellular reprogramming.¹⁴⁹ Given that dAC is able to 'reset' the epigenetic patterns of cells by erasing methylation, it has been successfully used to chemically induce reprogramming of somatic cells to pluripotent stem cells and neural cells (enhancing and in some cases replacing reprogramming transcription factors).¹⁵⁰

Azanucleosides as therapeutic agents

As previously mentioned, dAC has been clinically approved for the treatment of MDSs.¹²⁰ This is a group of diseases that affects normal blood cell production in the bone marrow. As a result, in MDS, the bone marrow produces abnormal, immature blood cells that fail to mature and work properly. At low concentration dose schedules, dAC has been reported to have significant activity against MDS.¹⁵¹ It has also shown activity against AML and Chronic Myeloid Leukaemia (CML). Several trials suggested that the effect of dAC could be further enhanced with combination treatments with growth factors, HDAC inhibitors and other chemotherapeutic agents such as cytarabine or gentuzumab.¹²⁰ Although azanucleosides are only used for treatment of myeloid disorders, a number of recent clinical trials indicate that dAC has potential against other cancer types such as ovarian, colorectal, cervical cancer and other solid tumours.¹⁵² It has also been indicated that dAC could enhance the sensitivity of tumour cells to treatments against which they previously acquired resistance.¹⁵³

Limitations of azanucleosides

Despite the promising efficacies and potential applications of azanucleosides, their use have several limitations. One of the main challenges of AC and dAC is their chemical instability in aqueous solutions due to the opening of the triazine nucleobase.^{154,155} A simple model of dAC's decomposition is shown in Figure 1.4. The triazine ring is prone to nucleophilic attack by water at position 6 (I), which could then undergo subsequent ring opening (II) and deformylation (III). The rate of degradation greatly depends on the pH and temperature of the solution.¹⁵⁵ Higher pH and higher temperatures increase the rate of degradation. The measured half-life of the drug is 10 hours at 37 °C in potassium phosphate buffer pH 7.4.¹⁵⁵



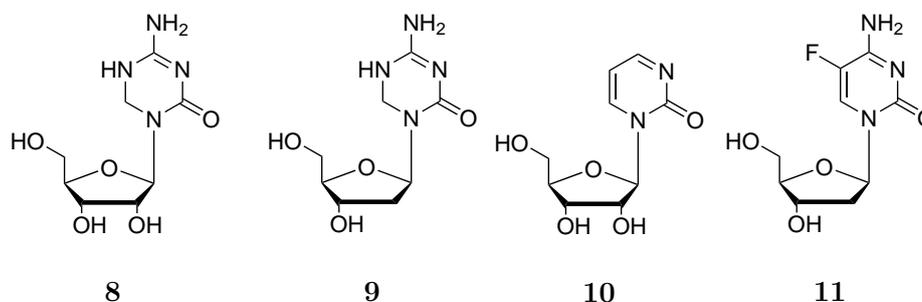
Scheme 1.4 Hydrolytic instability of dAC in aqueous solution

The clinical use of azanucleosides is also limited due to its cytotoxic side-effects. Due to the chemical instability and cytotoxicity of azanucleosides, prolonged administration at lower doses are preferable in the clinic. In addition, azanucleosides are rapidly metabolized by CDA following administration.¹⁵⁶

1.3.4 Chemical tools IIb: nucleoside inhibitor analogues

Alternative cytidine nucleoside analogues

To improve the stability and efficacy of azanucleosides, several alternative cytidine analogues have been developed in recent years (Scheme 1.5).^{103,157}



Scheme 1.5 Structures of second generation nucleoside DNMT inhibitors

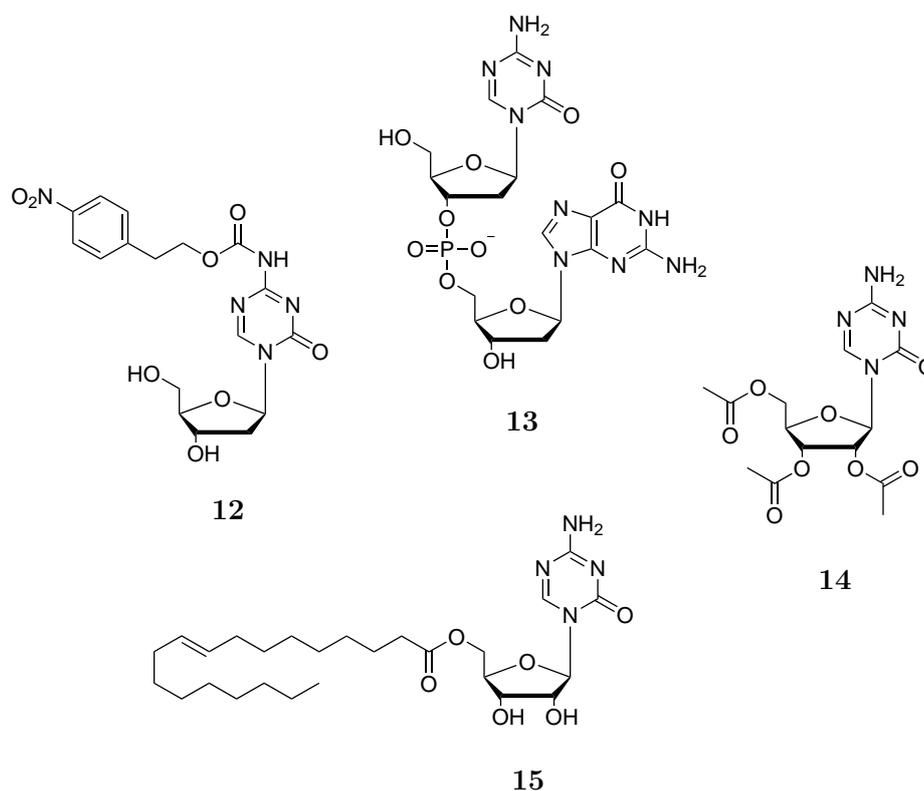
5,6-dihydro-azacytidine (**8**) and 2'-deoxy-5,6-dihydro-azacytidine (**9**) are hydrolytically stable analogues of AC and dAC.^{158,159} The saturation of the 5,6-double bond in the triazine ring abrogates nucleophilic attack of position 6 by water. Unfortunately DHAC has exhibited cardiac toxicity in clinical studies,¹⁵⁸ but DHDAC is well tolerated and has been shown to have promising activity as an anti-HIV agent.¹⁶⁰ Its prodrug KP1461, is currently under clinical investigation.¹⁶¹ Crystallographic studies have also indicated that **9** is able to inhibit DNMT1 by occupying the active site of the enzyme as a transition state mimic rather than forming a covalent bond like dAC.¹⁶⁰ Although **9** is efficient as a free nucleoside analogue to inhibit DNMT, its hypomethylating potency is much lower than that of dAC.¹⁵⁹

Zebularine (**10**) is a cytidine analogue that lacks an amino group at position 4 of the pyrimidine ring. Its greatest advantage is its stability against cytidine deaminase, which enabled this drug to be taken orally.¹⁶² However, inefficient metabolic activation and unexpected fatal toxicity found in monkeys hampered further clinical development of this agent.⁹¹

The next cytidine analogue, 5-fluoro-2'-deoxycytidine (**11**) has also shown promising hypomethylating activity in human breast and lung carcinoma cells.¹⁶³ Compound **11** is stable in aqueous solution, however its rapid metabolism by cytidine deaminase remained an issue. In addition, its metabolic side-product, 5-fluoro-2'-deoxyuridylate turned out to be an inhibitor of thymidilate synthase which limited its further clinical development.¹⁶⁴

Prodrugs of azanucleosides

In addition to exploring alternative cytidine analogues, a number of groups have focused on the development of prodrugs for AC and dAC in order to improve their stabilities and cellular delivery (Scheme 1.6).



Scheme 1.6 Structures of azanucleoside prodrugs

Güimil and co-workers synthesised 4-NPEOC-dAC (**12**), which contains a 4-nitrophenethyl carboxy (4-NPEOC) group at position 4 of the triazine ring.¹⁶⁵ This group protects dAC from degradation by CDA enzyme and could therefore prolong the lifetime of dAC. Preclinical studies have shown that the 4-NPEOC group could be selectively removed

by human carboxylesterase 1 in liver cancer cell lines. However, 4-NPEOC-dAC **12** is less potent than dAC and its dependence on the carboxylesterases would limit its potential activity to organ with high concentration of this enzyme, such as the liver.¹⁶⁶

Astex Pharmaceuticals have recently developed SGI-110 (**13**), a dinucleotide prodrug of dAC. In SGI-110, the active dAC is linked to deoxyguanosine (dG) *via* a phosphodiester bond. This linkage results in reduced susceptibility to degradation by CDA and provides prolonged plasma stability and high tolerance in humans.¹⁶⁷ Since 2010, a number of clinical studies began to evaluate the safety and efficacy of SGI-110 in various haematological and ovarian cancers.¹⁶⁸

Efforts were made not only to improve the stability of azanucleosides, but also to improve their delivery within the cell and the entire body. CP-4200 (**15**) is an elaidic ester analogue of AC developed in order to render AC less dependent on conventional nucleoside transport systems.¹⁶⁹ This analogue has shown superior efficacy to **5** in mouse tumour models. In addition, to improve oral bioavailability, 2',3',5'-triacetyl-5-azacytidine (**14**) was developed as a potential prodrug for AC.¹⁷⁰ The prodrug demonstrated significant pharmacokinetic improvements in bioavailability, solubility and stability over AC. Although this analogue was not investigated further, different oral formulations of AC (CC-486) have also been developed and are currently under clinical investigations.¹⁷¹ An orally administrable version of dAC involving its mesylate salt was also reported in a preclinical study.¹⁷²

1.3.5 Need for controlled DNA methylation

In addition to the ever growing interest in its therapeutic effects in clinical settings, dAC continues to serve as an immensely valuable research tool in the field of epigenetics and developmental biology. It remains the most efficacious demethylator despite the numerous attempts to develop alternative nucleosides with improved stability profiles. Due to its challenging chemistry, only a few dAC prodrugs were successfully made with reported improvements in their pharmacological properties such as their uptake¹⁷³ and reduced toxicity.^{166,174} In most cases, intracellular enzymes (such as esterases and phosphatases) serve as 'activators' of these prodrugs. It remains a challenge to control the intracellular levels of these enzymes and thereby the activity of these dAC analogues.

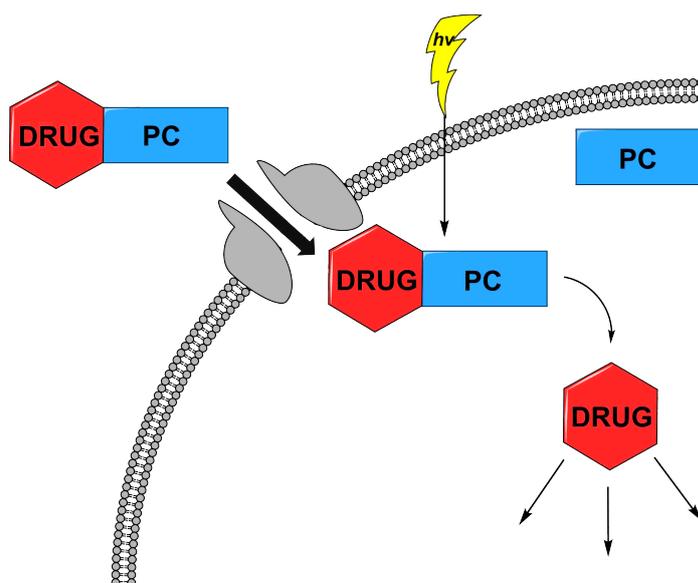
Currently there is no small molecule with a tuneable bioactivity that can modulate DNA methylation. A molecule with a tuneable bioactivity would enable the study of methylation changes in the same biological environment, thereby further aiding our understanding of the functional significance of DNA methylation as part of many biological processes. The therapeutic advantages of tuneable DNA demethylators would be the potential targeted approach of treatments with reduced cytotoxic side effects. Since azanucleosides have high demethylating efficacy and potential to reactivate tumour suppressor genes, the possibility to develop analogues with tuneable bioactivity would be very desirable.

Light is an excellent tool that could provide control over the bioactivity of small molecules in a spatiotemporal confinement.¹⁷⁵ Optical control of histone acetylation via small molecules equipped with photoswitches has been recently reported.¹⁷⁶ In addition, light-tuneable versions of the TALE and CRISPR-Cas systems have also been explored to edit the epigenome with moderate to good efficiencies.^{177–180} However, no chemical tools have been developed to date that could target DNA methylation in a light-dependent manner. A light-tunable methylation tool would enable researchers to direct methylation changes to specific areas in the cell or tissues. In order to make light-tuneable small molecules, one could pursue a photocaging strategy.¹⁸¹ Photocaged compounds are light-sensitive probes that functionally encapsulate biomolecules in an inactive form.¹⁸²

1.4 Tuning bioactivity using light: Photocaging

1.4.1 Principles of photocaging

The principal idea behind photocaging is that a biologically active molecule of interest can be rendered temporarily inactive by chemical modification or 'caging' with a photolabile protecting group (also termed as phototag or photocage).¹⁸¹ The 'caged' compound can enter the cell via similar transport systems as their active equivalent. (Figure 1.7) Restoration of biological activity can be achieved by irradiation of the caged compound at a particular wavelength.¹⁷⁵ This results in a photochemical cleavage of the phototag and simultaneous concentration spike of the unmodified active substance, which can then exert its biological effect.¹⁸² This process is also termed as 'uncaging', 'photouncaging' or 'deprotection'.



Scheme 1.7 Light activation of a photocaged-compound

Light is an ideal and orthogonal signal.¹⁷⁵ It easily penetrates cell membrane, and provided the wavelengths used are not too short, cells are not harmed by light. Furthermore, light can achieve both global and focused illumination on cells, depending on the light source and techniques used.¹⁸² Light cannot only be directed, but it can also be modulated in time and amplitude. Thus, uncaging can be used to produce rapid, repetitive release of biomolecules and it can also be finely tuned for gradual uncaging of the compound. Overall, these properties of light triggered uncaging would enable one to control the location, time and dosage of a drug once it has entered the cell.

1.4.2 Phototags

The choice of the phototag plays a crucial role in the behaviour of the caged compound and the properties of its uncaging. In the first part of this section, a number of important characteristics of phototags will be discussed, followed by a brief review of the two most popular phototag families used in literature.

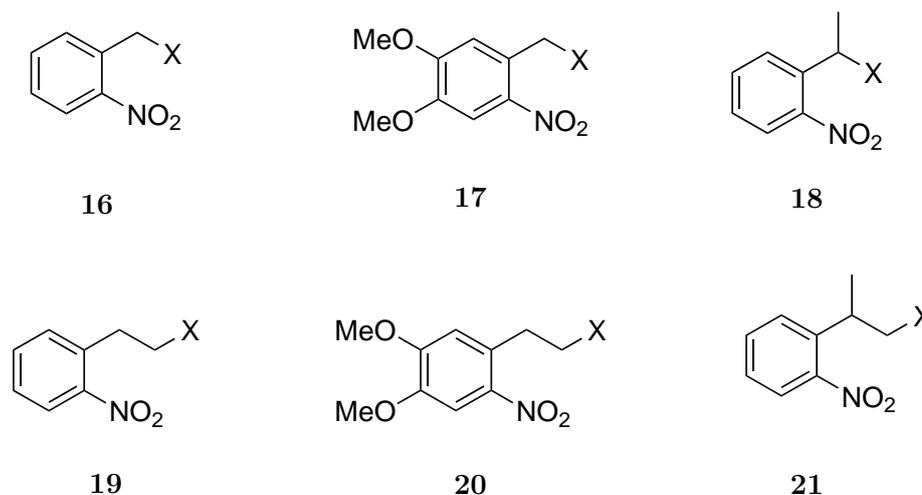
Important characteristics of phototags

The criteria and design of a good phototag depends on its application. For cell biological applications, the phototag should have strong absorption (ϵ) at wavelengths that are less damaging to the biological environment ($\lambda_{max} > 300$ nm).¹⁸¹ A good phototag should also have high aqueous solubility and be able to pass through biological barriers such as cell membranes. Furthermore, the photochemical deprotection ('uncaging') of the caged compound should be complete: it should only release the 'uncaged' product and the phototag. The efficiency of release is another important factor. The quantum yield of uncaging (Φ) is the measure of the efficiency with which the light absorbed converts the caged compound to its uncaged derivative.¹⁸¹ It is the number of molecules converted per number of photons absorbed. The closer this value is to 1, the more efficient is the uncaging process. And finally, when the photouncaging is complete, the remaining byproducts should be transparent and non-toxic to the cells. Several phototags meet many of the criteria above.

Nitroaryl-based (*o*-NB and *o*-NPE) phototags

The *ortho*-nitrobenzyl (NB) and *ortho*-nitrophenethyl (NPE) compounds are the most commonly used phototags to date. Their popularity can be attributed to their commercial availability and the ease by which they can be conjugated to bioactive molecules. Structures of the frequently used derivatives are shown in Scheme 1.8.

The first *o*-nitrobenzyl- (NB) analogue **16** was reported as a general photolabile protecting group in 1970.¹⁸³ Its derivatives carry a leaving group at the benzylic position to release the protected substrate upon irradiation. Substrates can include alcohols, amines, but carbonic acid derivatives such as carbonates and carbamates are the most convenient choices to synthesise due to the ease of chemical conjugation.¹⁸⁴

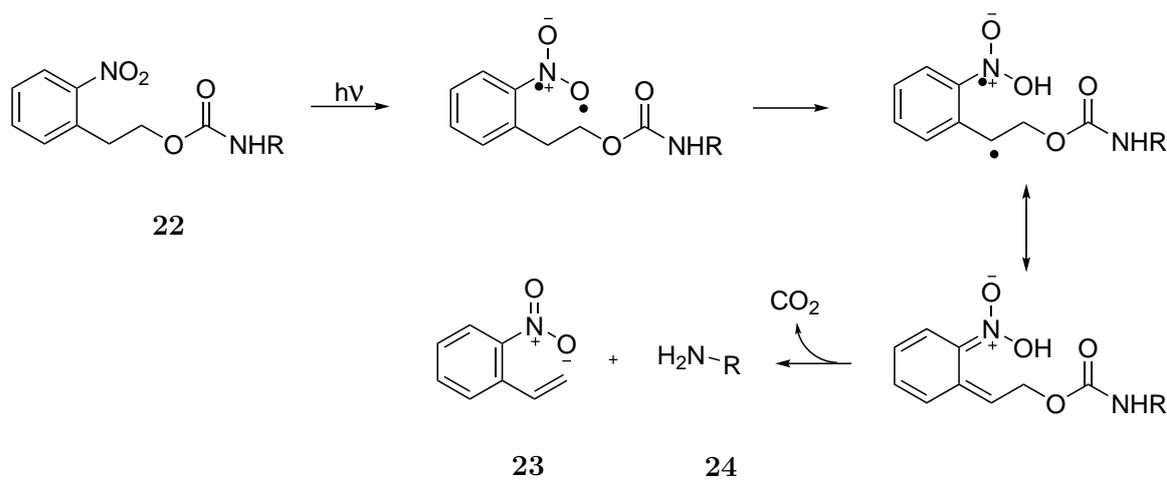


Scheme 1.8 Popular NB and NPE phototags in literature

Modifications in the aromatic ring can greatly affect absorbance and solubility of NB phototags. For example, the addition of two methoxy groups in 4,5-dimethoxy-2-nitrobenzyl (DMNB) **17** resulted in a 95 nm increase in λ_{max} . Furthermore, substituents in the benzylic position (such as a methyl group in methyl-2-nitrobenzyl **18** can also positively affect the quantum yield of uncaging (from $\Phi = 0.13$ to $\Phi = 0.64$).¹⁸⁵ Other groups such as the -cyano, -carboxylic acid and other NB rings were also explored, but their effect on quantum yield was not as great as for the methyl group.¹⁸¹

While the photochemical properties of NB analogues were extensively studied, the main drawback of these phototags remained the toxic nitrosoaldehyde side-products that were formed upon photolysis.¹⁸⁶ This issue was overcome by the NPE phototags, which release a more biocompatible nitrostyrene byproduct. These compounds were discovered in the late 1990s by Hasan and coworkers.¹⁸⁷ In addition to the improved biocompatibility, the first NPE phototag **19** had a greater quantum yield of uncaging ($\Phi = 0.042$) compared to its NB equivalent **16** ($\Phi=0.033$).¹⁸¹ Similarly to the NB phototags, the addition of two methoxy groups in 4,5-dimethoxy-2-nitrophenethyl phototags **20** improved the absorption maximum by nearly 100 nm. Analogue **21** had eight times higher quantum yield ($\Phi = 0.35$) compared to the parent NPE phototag **19**.¹⁸⁸

Despite the similarity in their structures, the NB and NPE phototags undergo markedly different uncaging mechanisms.¹⁸¹ The mechanism of uncaging of a NPE-protected carbamate **22** is shown in Scheme 1.9. Firstly, light irradiation induces radical formation within the nitro group. These radicals proceed with an intramolecular hydrogen abstraction at the α -position, which is followed by β -elimination of the protected substrate. This gives the nitrostyrene byproduct **23** with concomitant liberation of CO_2 and the unprotected amine **24**.



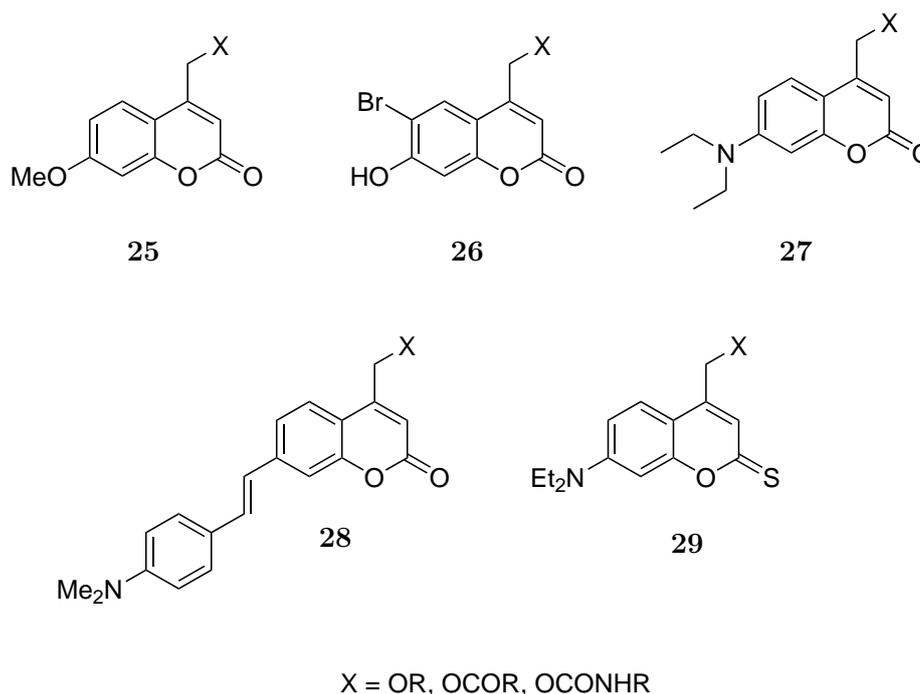
Scheme 1.9 Photodeprotection of an NPE-derived carbamates¹⁸¹

In summary, NPE compounds are easily accessible phototags with excellent photochemical characteristics. They also serve as good reference compounds for photouncaging. However, their use is more suitable with chemical applications due to their limited aqueous solubilities. Notable applications of NPE phototags include light-automated DNA oligonucleotide synthesis¹⁸⁸ and transient disruption of DNA hybridization.¹⁸⁹

Coumarin-based phototags

The coumarin compounds are attractive phototags for cellular applications not only because they have improved aqueous solubility profiles, but they exhibit better photouncaging properties at longer wavelengths (up to $\lambda = 450$ nm).¹⁸¹ They have very high molar absorption coefficients and high quantum yields and as a result, faster release rates at longer wavelengths compared to their equivalent NB/NPE caged compounds.¹⁸⁶

The first analogue of the coumarin-based phototags, 7-methoxycoumarin **25** ($\lambda_{max} = 320$ nm) was used to cage diethyl phosphate derivatives and was reported by Givens *et*

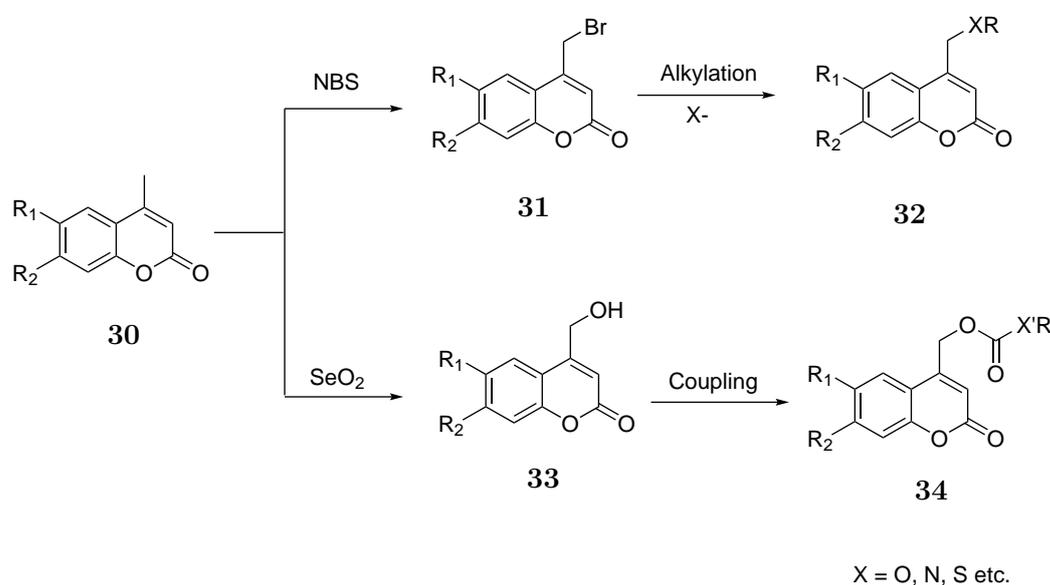
Scheme 1.10 Structures of coumarin phototags¹⁸¹

*al.*¹⁹⁰ Modifications in the C6 and C7 position enabled further improvement in water solubility profiles and increase in absorption maxima. Most notably in the 6-bromo-7-hydroxycoumarin phototag **26** (BHC), the bromine substitution was designed to lower the pH of the $-OH$ group at position 7 by two units to achieve its complete deprotonation at physiological pH.¹⁹¹ This enhanced the aqueous solubility and in addition increased maximum wavelength of absorption by 50 nm ($\lambda_{max} = 370$ nm). But so far, the most successful modification to date is the introduction of a 7-amino substituent in 7-diethylaminomethylcoumarin (DEACM) **27**. With the highest quantum yield among all the coumarin analogues, DEACM has recorded quantum yields between $\Phi = 0.21$ and $\Phi = 0.28$ and it also has the higher absorption maxima ($\lambda_{max} = 390$ nm).¹⁹²

Further efforts were made to push the (λ_{max}) to even higher wavelengths, towards $\lambda_{max} = 500$ nm. Bao *et al.* introduced an additional styryl group at position 7 giving styrylcoumarin **28** to increase the absorption wavelength to $\lambda_{max} = 407$ nm.¹⁹³ Substitution of the carbonyl group of the coumarin ring to a thionyl group in analogue **29** also resulted in a further bathochromic shift of 100 nm, reaching a final $\lambda_{max} = 487$ nm.¹⁹⁴ Although these agents have compromised quantum yields at high wavelengths

($\Phi_{(420nm)} = 0.00079$ and $\Phi_{(470nm)} = 0.005$), the prospect of long wavelength phototags are promising as they allow deeper penetration and reduced toxicity by UV light.

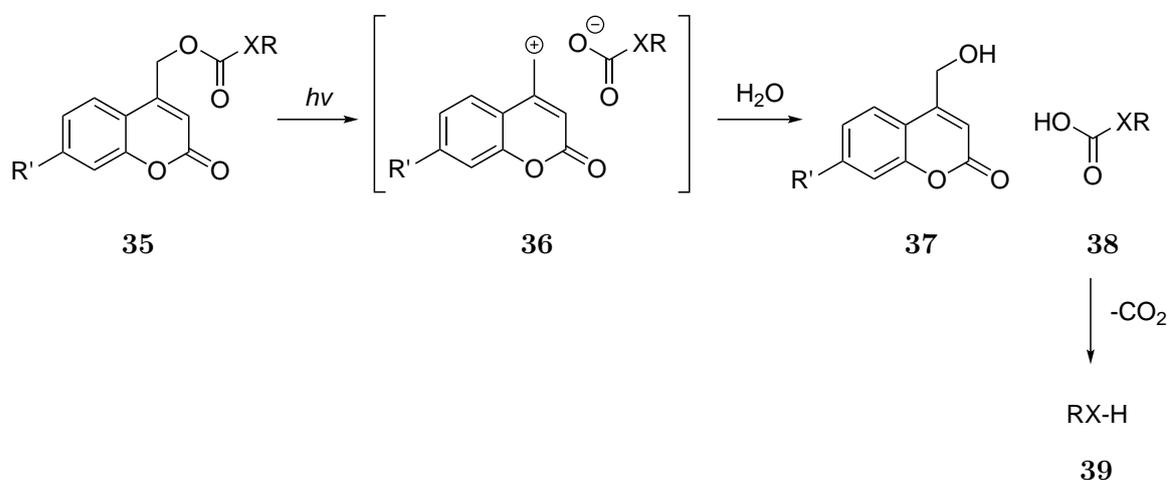
The synthetic strategies towards coumarin-caged compounds are summarised in Scheme 1.11. Substrates are generally attached to the 4-methyl position of the coumarin scaffold. The desired methylcoumarin derivatives **30** can be commercially purchased or synthesised *via* Pechmann condensation.^{195,196} The methylcoumarin can either be converted into a halide derivative such as 4-bromo-coumarin **31** *via* treatment with *N*-bromosuccinimide (NBS) or a hydroxy derivative (coumarinyl-4-methanol, **33**) *via* oxidation with selenium dioxide SeO₂.¹⁸¹ The bromide handle can provide access to thiols, alcohols, amines *via* direct alkylation to give **32** or carboxylic acids and amino acids *via* S_N2 substitution. Conversely, the hydroxycoumarin **33** derivative is a common precursor to carbonates, carbamates, esters, phosphates and sulphonates **34**.¹⁸¹



Scheme 1.11 Synthetic strategies towards coumarin-caged compounds

The suggested uncaging mechanism of coumarin-caged compounds containing a carboxy linkage is shown in Scheme 1.12. The first step upon light irradiation is the heterolytic (C–O) bond cleavage within the coumarin-caged compound **35**. This is followed by a formation of a tight ion pair **36**, which then reacts with water to form a stable coumarinmethyl alcohol **37** and the unstable liberated carboxylate residue **38** undergoes decarboxylation to give the uncaged product **39** as a free alcohol or amine. The heterolytic bond cleavage is thought to proceed rapidly, whereas ion recombination is slower and decarboxylation is the slowest, rate-limiting step. Alcohols, phenols, thiols and amines are poor leaving groups, while carbamates and carbonates tend to be released faster and more efficiently. In the case of a coumarin-caged compound

with no carboxyl linkage, the mechanism is identical with the omission of the final decarboxylation stage and the direct release of the uncaged substrate.



Scheme 1.12 Suggested deprotection mechanism of coumarin-caged molecules

DEACM **27** is one of the most commonly used coumarin phototags in literature. This is indicated from a number of studies where DEACM phototag was used to cage a variety of biomolecules, for example Hagen and co-workers photocaged cyclic nucleotide analogues to selectively activate cyclic nucleotide-gated ion channels in human embryonic kidney (HEK293T) cells.¹⁹² In addition, DEACM was also used to cage adenosine triphosphate (ATP) to control transcription *in vitro* using light in a study by Pinheiro *et al.*¹⁹⁷ Furthermore, DEACM-caged cholesterol was also synthesised for the purpose of studying sterol pathways.¹⁹⁸

1.4.3 Photocaging nucleosides

In this project, we are interested in photocaging azanucleosides. Given that they are nucleoside analogues, the past caging strategies for nucleosides will be reviewed in this section. Nucleosides can be caged at three different positions: 5'-position, 3'-position and the nucleobase (Figure 1.10). For each position, the general synthetic strategies and applications of the resulting caged-nucleosides will be discussed.

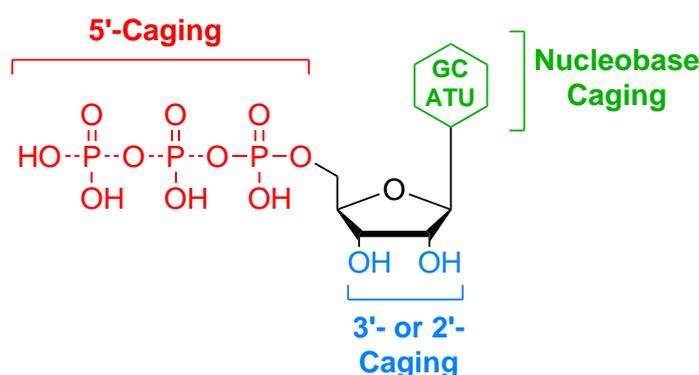


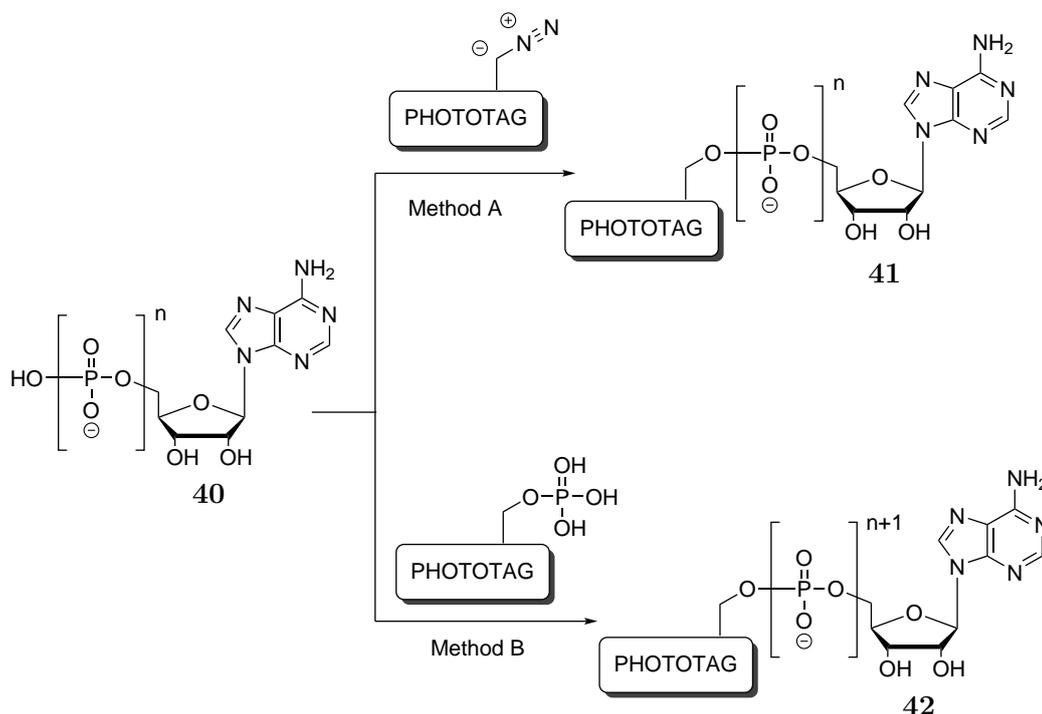
Fig. 1.10 Summary of nucleoside caging strategies

Caging the 5'-position

Caged-ATP was the first protected nucleoside reported by Kaplan and coworkers.¹⁹⁹ In their study, the NB phototag was linked to the γ -phosphate group on ATP. Caged-ATP enabled researchers to analyse fast, energy-requiring biological events such as ion transport and signal transduction. Both NB and coumarin photocages have been used to cage ATP, GTP and other cyclic nucleotides.^{197,200,201}

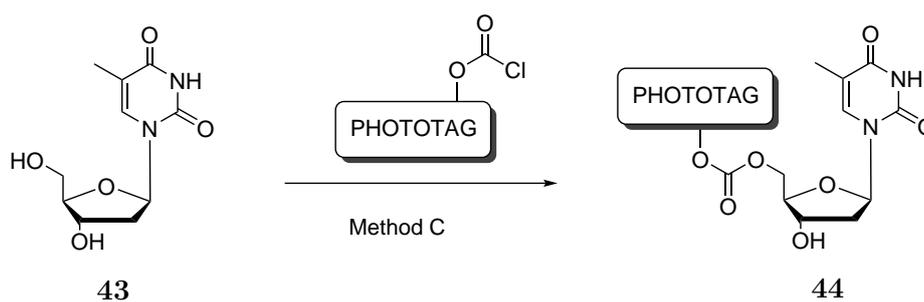
The majority of the studies focused on the photocaging 5'-phosphorylated nucleosides. Historically, there are two main methods by which the phototag could be linked to this position (Scheme 1.13). Method A involves the direct coupling of the diazo derivative of the phototag to the nucleoside mono-, di- or triphosphate **40**. Geissler and *et al.* explored this method with coumarin phototags, but their reported yields to obtain **41** were very low (1.7%) due to the low reactivity of 4-(diazomethyl)coumarins.²⁰¹

Conversely, method B is significantly higher yielding despite requiring multiple synthetic steps. The phototag phosphate is prepared from the phototag alcohol in 3 steps (not shown) *via* treatment with a phosphoramidite, oxidation with tert-butyl hydroperoxide and subsequent hydrolysis with trifluoroacetic acid.²⁰¹ The phototag phosphate is then



Scheme 1.13 Nucleoside photocaging strategies at 5'-position I.

coupled with carbonyldiimidazole-activated nucleoside monophosphate or diphosphate to give analogue **42**.²⁰² In addition to photocaging of the phosphate group, direct caging of the 5'-OH of nucleoside **43** *via* phototag-chloroformate has also been reported by Bühler *et al* (Scheme 1.14).²⁰³ In the resulting photocaged-nucleoside **44**, 5'-OH and the phototag was linked *via* a carbonate bond.



Scheme 1.14 Nucleoside photocaging strategies at 5'-position II.

Caging the 3'- and 2'-position

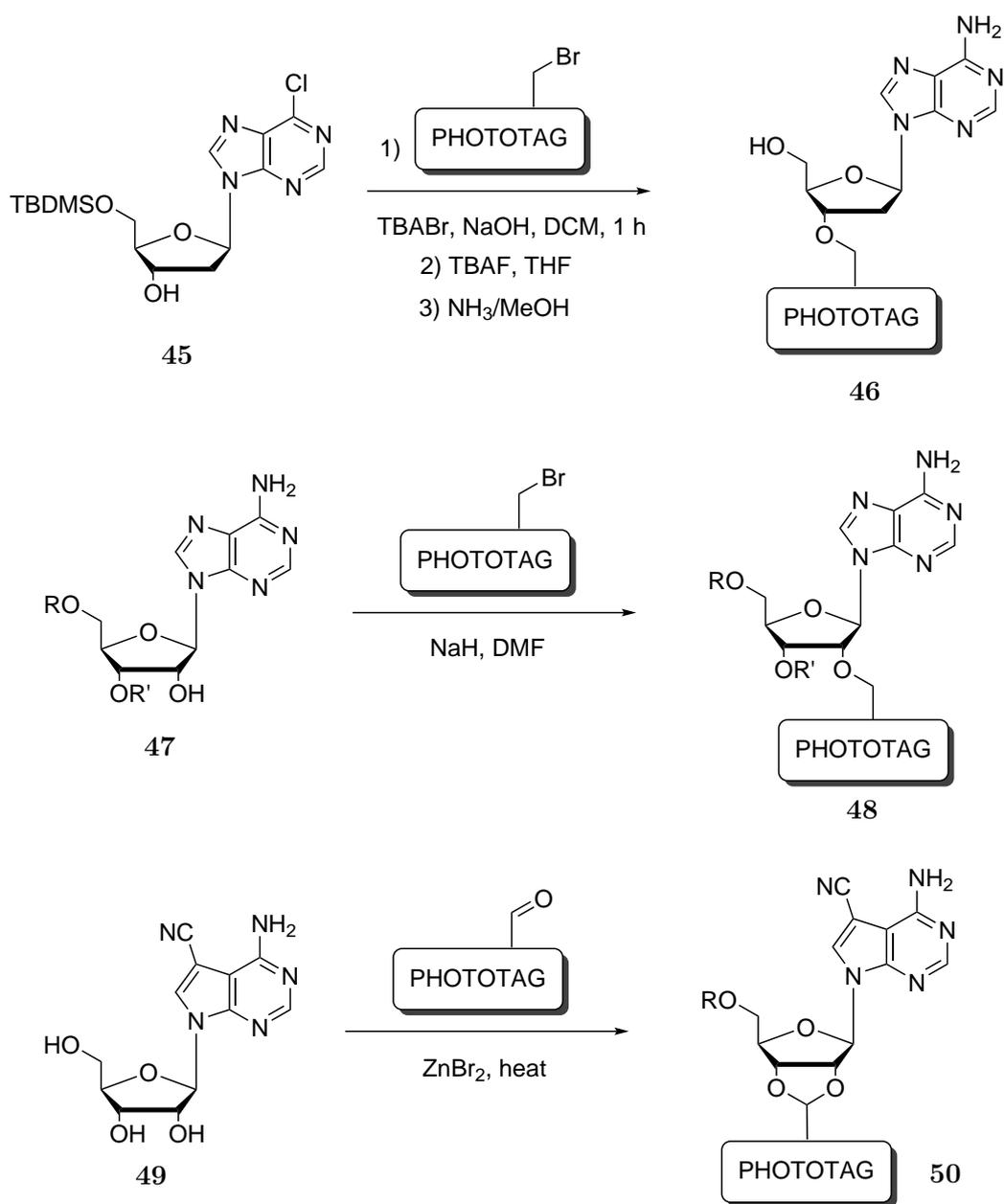
The 2'- and 3'-OH groups are important positions for biomolecular recognition.²⁰⁴ The 3'-OH is essential for chain elongation during DNA synthesis. Thus, temporarily blocking this position terminates the polymerase chain reaction and prevents further incorporation of nucleotides into DNA. Photocaging the 3'-OH creates reversible chain terminators, which were particularly useful for improving pyrosequencing techniques.²⁰⁵ Photocaging of the 2'-OH ribonucleotides enabled researchers to study the mechanism of RNAzyme in detail.²⁰⁶

Synthetic strategies for photocaging these positions are summarised in Scheme 1.15. 3'-caging so far has only been reported *via* direct alkylation of the 3'-OH by reacting starting nucleoside **45** with the phototag-bromide in the presence of NaOH to give photocaged analogue.²⁰⁷ It is essential that the 5'-OH and the nucleobase are protected prior to the photocaging.

In this example, the 5'-OH is protected with a *tert*-butyldimethylsilyl (TBDMS) group and the exocyclic amine of adenine nucleobase is replaced with a Cl group. Direct coupling to the photocage-bromide gives an ether-linked photocage at the 3'-position. The removal of the protecting groups is achieved by treatment with tetrabutylammonium fluoride (TBAF) followed by methanolic ammonia (NH₃/MeOH) to give the desired 3'-protected nucleoside analogue **46**.

The functionalisation of 2'-OH like in analogue **47** is similarly carried out in the presence of base (NaH), and it requires previous protection of 3'- and 5'-OH groups.²⁰⁸

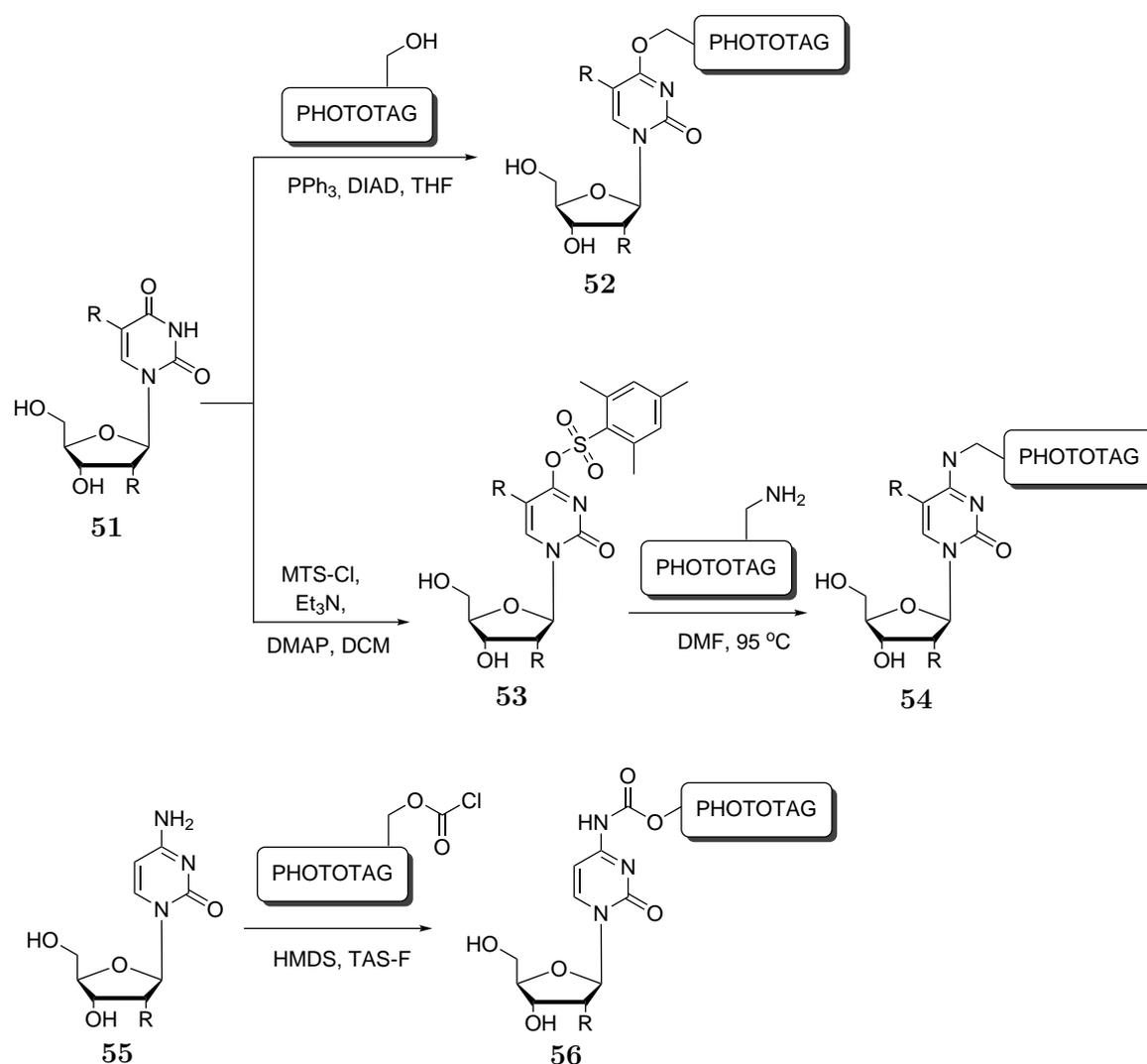
The final photocaged analogue **48** is linked *via* an ether linkage in position 2. Simultaneous protection of 2'- and 3'-OH groups is also possible. Deiters *et al.* reported the photocaged version of an adenosine analogue Toyocamycin **49**. This analogue can efficiently inhibit ribozyme function. The photocaging of Toyocamycin was achieved in the presence of photocage-aldehyde derivative and ZnBr₂ and heat to give caged analogue **50**.²⁰⁹



Scheme 1.15 Nucleoside photocaging strategies at 3'-position

Caging the nucleobase

Placing the phototag on the nucleobase transiently blocks base-pairing properties and enzymatic recognition.²⁰⁴ Due to these properties, nucleobase caging has been extensively used in molecular biology - to control the formation of DNA structures upon light exposure.²¹⁰ Other applications include light regulation of transcriptional activity,²¹¹ secondary and tertiary structure formation,²¹² DNAzyme activity²¹³ and even DNA replication.²¹⁴



Scheme 1.16 Nucleoside photocaging strategies at the nucleobase

Photocages are generally linked to the exocyclic amine (cytidine, adenine) or the carbonyl moiety (thymidine and guanosine) of nucleobases. For simplicity, only caging of cytidine and thymidine nucleobases will be presented here. Caging strategies are summarised in Scheme 1.16. Starting from the thymidine/carbonylated dC **51**, one can couple the phototag-OH *via* a Mitsunobu-like reaction in the presence of triphenylphosphine (PPh₃) and diisopropyl azodicarboxylate (DIAD). In the resulting photocaged analogue **52**, the phototag is linked *via* an ether bond.

In order to photocage cytosine nucleobases *via* an amine linker, one would also start from the carbonylated version of the nucleobase **51**. But in this case, activation of the C-O position is necessary with a 2-mesitylenesulfonyl-group. Addition of mesitylsulphonyl-chloride (MTS-Cl) to **51** gives the activated **52**. This activating group can be easily replaced when the amine-photocage derivative is added. As a result, the phototag is successfully installed on the nucleobase *via* an amine linker (**54**).

Furuta and coworkers reported a milder caging strategy that involves the direct coupling of the photocage-chloroformate to cytidine **55**.²¹⁵ In this case, it is essential that the 3'- and 5'-OH groups are transiently protected to selectively target the nucleobase. In the resulting photocaged-analogue **56**, the phototag is linked to cytidine *via* a carbamate bond.

1.5 Aims of the project

In this project, we aimed to pursue the photocaging strategy by designing and synthesising photo-responsive ‘caged’ derivatives of azanucleosides that can achieve tissue-specific hypomethylation under spatiotemporal control. (Figure 1.11)

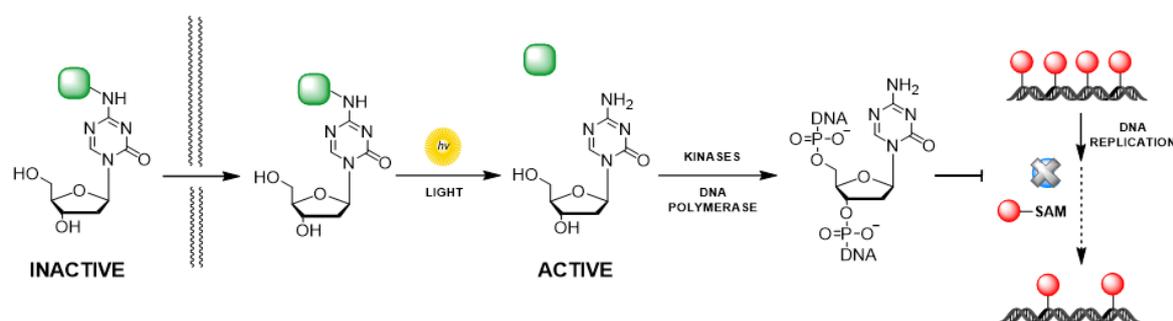


Fig. 1.11 Aims of the project

There were three main aims for this project:

- (1) Chemical design and synthesis of caged-azanucleoside analogues
- (2) Photochemical characterisation and stability analysis of these analogues
- (3) Cell biological studies to inspect light-induced demethylating activity *in vitro*

It is envisaged that photocaged azanucleosides would be able to enter the cell through the same nucleoside transport system utilized by azanucleosides. Depending on where the phototag is placed, it is assumed that the phototag would prevent the formation of key interactions (dCK, polymerization or phosphorylation). Restoration of biological activity can be achieved by irradiation of the caged compound (uncaging), resulting in bioactivity only at the site of illumination, which would greatly reduce unwanted cytotoxicity of healthy cells. If successful, the photocaged azanucleosides would have the same clinical benefits as the original azanucleosides but the cytotoxic side-effects would be greatly mitigated.

Photocaged azanucleosides may also hold great promises as research tools. As light tuneable demethylators, they would be able to induce methylation changes in a tissue-specific manner. This would be invaluable in the field of developmental biology and regenerative medicine where the differential effects of methylation changes could be compared in the same environment.

Chapter 2

Synthesis of Photocaged Azanucleosides

2.1 Introduction

Photocaging azanucleosides offers light-tuneable, spatiotemporal control over their bioactivity. Similarly to conventional nucleosides, the activity of azanucleosides can be blocked by modification at the nucleobase, 3'- or 5'-OH groups. A phototag on the exocyclic amine of the nucleobase or the 5'-OH prevents the phosphorylation of azanucleosides by dCK, while placing the phototag on the 3'-terminal blocks chain elongation.

In this chapter, the design and synthesis of caged azanucleoside analogues will be discussed. NPE and coumarin-derived phototags were attached to all three possible caging positions of azanucleosides. The photocaged analogues were tested to examine the influence of the phototag and the caging-position on the stabilities and photochemical properties of the compounds.

2.2 First generation photocaged azanucleosides

Given the abundant literature on nucleobase modifications, the exocyclic amine was chosen as the first site for photocaging. It was also hypothesised that a phototag on the exocyclic amine could protect azanucleosides from degradation by cytidine deaminase.

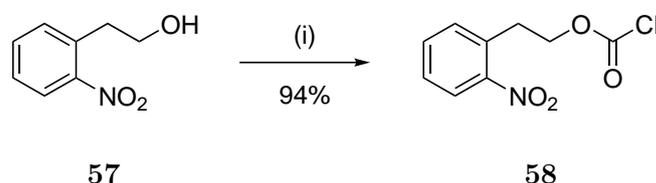
Previous work in the Howorka group attempted to photocage 5'-azacytidine with NB-derived phototags, but unsuccessfully. Coupling of NPE-phototags were successful, but only when added in great excess. Here, we describe the preparation of NPE-derived phototags followed by the successful coupling strategy that was adapted from a protocol by Güimil *et al.*¹⁶⁵

2.2.1 Design and synthesis

Synthesis of NPE-caged azanucleosides

NPE-derived phototags are popular due to their easily accessible structures and good photouncaging properties.¹⁸¹ Our aim was to attach the NPE phototags to azanucleosides *via* carbonate or carbamate linkers as these derivatives have reportedly better uncaging properties compared to their ether or amine linked counterparts.

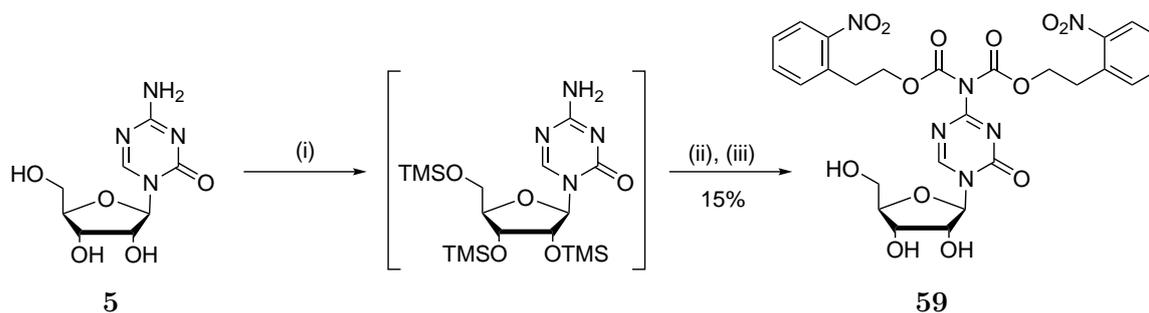
The first step was to convert phototags bearing hydroxyl groups into reactive chloroformates. The commercially available 2-nitrophenethyl alcohol **57** was mixed with phosgene overnight to afford the corresponding chloroformate **58** in 94% yield (Scheme 2.1).



Scheme 2.1 Synthesis of NPE-chloroformate **58**. *Reagents and Conditions:* (i) COCl₂, THF, 0°C, 1 h then rt, 16 h

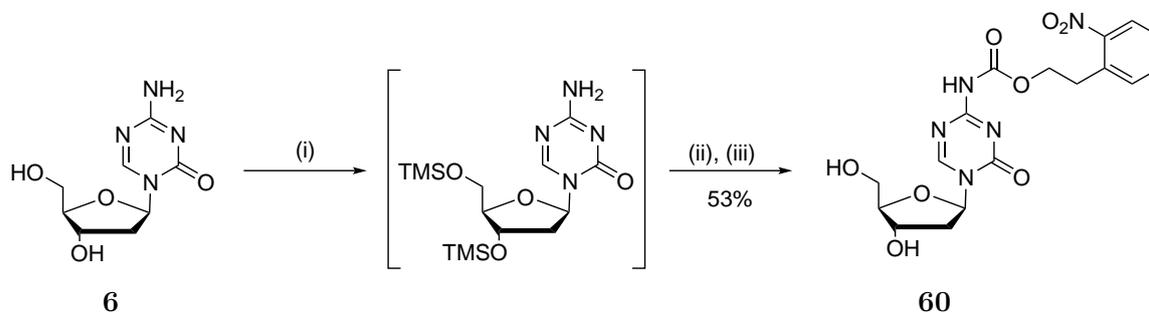
The resulting chloroformate **58** was then coupled to 5-azacytidine (AC) **5** *via* a one-pot transient protection strategy (Scheme 2.2).¹⁶⁵ AC was first treated with hexamethyldisilazane (HMDS) in order to protect the 3'- and 5'- hydroxyls groups from unwanted side reactions. This was followed by the addition of 11 equivalents of chloroformate **58** *in situ*, and the subsequent removal of hydroxyl protecting groups using tris(dimethylamino) sulfonium difluorotrimethylsilicate (TAS-F), an anhydrous source of fluoride ions.²¹⁶

Due to the instability of the triazine nucleobase, maintenance of anhydrous conditions was essential.¹⁵⁴ As a result, the first photocaged azanucleoside analogue **59** was resynthesised in 14.8% yield. However, this analogue had very poor aqueous solubility.



Scheme 2.2 Synthesis of bis-NPEOC-AC **59**. *Reagents and Conditions:* (i) HMDS, pyridine, DMF, rt, 2h; (ii) 11 eq chloroformate **58**, CH₂Cl₂/DMF, rt, 16 h; (iii) TAS-F, DMF, rt, 2 h

To address the solubility issue, the next aim was to attach a single phototag on the exocyclic amine using dAC **6** instead of AC as our substrate of interest. Attaching a single phototag would potentially improve the solubility due to the additional proton on the carbamate linker that could form hydrogen bonds with water molecules. Following the established protocol, two equivalents of chloroformate **58** were introduced to the silyl-protected dAC, affording mono-NPEOC-dAC **60** in 53% yield (Scheme 2.3). This analogue had slightly improved, but still modest aqueous solubility.

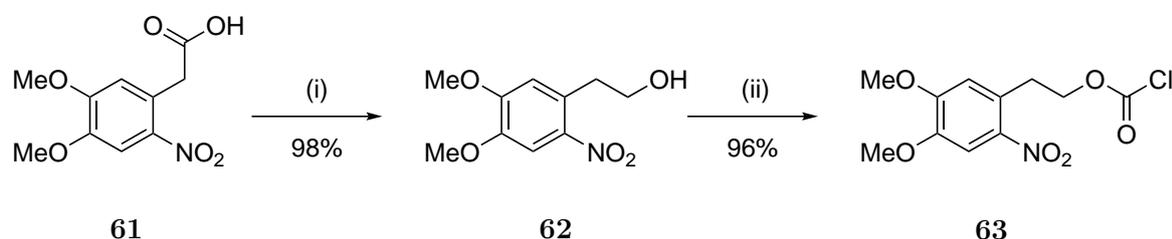


Scheme 2.3 Synthesis of mono-NPEOC-dAC **60**. *Reagents and Conditions:* (i) HMDS, pyridine, DMF, rt, 2h; (ii) 2 eq chloroformate **58**, CH₂Cl₂/DMF, rt, 16 h; (iii) TAS-F, DMF, rt, 2 h

Synthesis of DMNPE-caged azanucleoside

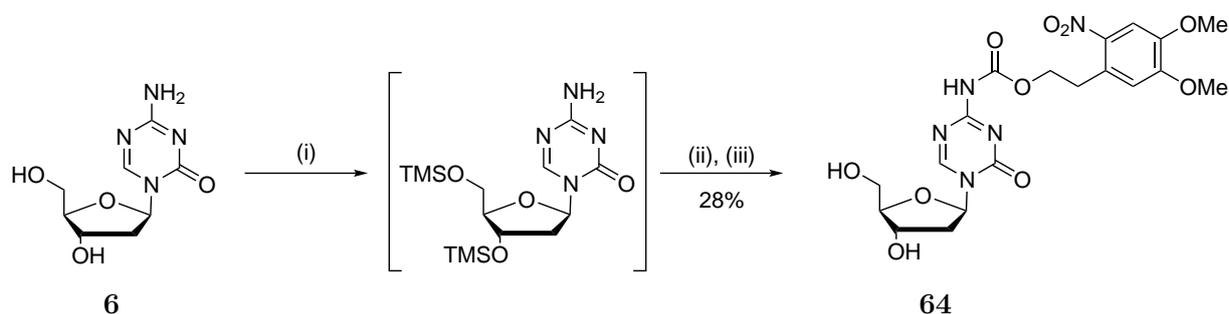
The 4,5-dimethoxy-2-nitrophenethyl (DMNPE) phototags were also used to photocage azanucleosides. These phototags have higher wavelength of maximum absorbance (λ_{max}) than their NPE equivalents due to the additional methoxy groups on their aromatic ring.

The synthesis of DMNPE phototags was achieved in two steps. Reduction of commercially available 4,5-dimethoxy-nitrophenylacetic acid **61** via BH_3 gave the corresponding alcohol **62** in 98% yield. This analogue was further subjected to overnight phosgenation to afford chloroformate **63** in 96% yield (Scheme 2.4).



Scheme 2.4 Synthesis of DMNPE-chloroformate **63**. *Reagents and Conditions:* (i) BH_3 , THF, rt, 2 h; (ii) COCl_2 , THF, 0°C , 1 h then rt, 16 h

Mono-DMNPE-dAC **64** was prepared at 28% yield *via* the established transient protection protocol and the addition of two equivalents of chloroformate **63** phototag (Scheme 2.5). In general, the synthesis of mono-tagged derivatives was achieved at significantly higher yields and this could be attributed to the greater hydrolytic stability of dAC compared to AC.



Scheme 2.5 Synthesis of mono-DMNPEOC-dAC **65**. *Reagents and Conditions:* (i) HMDS, pyridine, DMF rt, 2h; (ii) 2 eq chloroformate **63**, $\text{CH}_2\text{Cl}_2/\text{DMF}$, rt, 16 h; (iii) TAS-F, DMF, rt, 2 h

2.2.2 Photochemical characterisation

UV-Vis absorption

The UV-Vis absorption spectra of caged analogues and their components are depicted in Figure 2.1. The corresponding λ_{max} and ϵ values are summarised in Table 2.1.

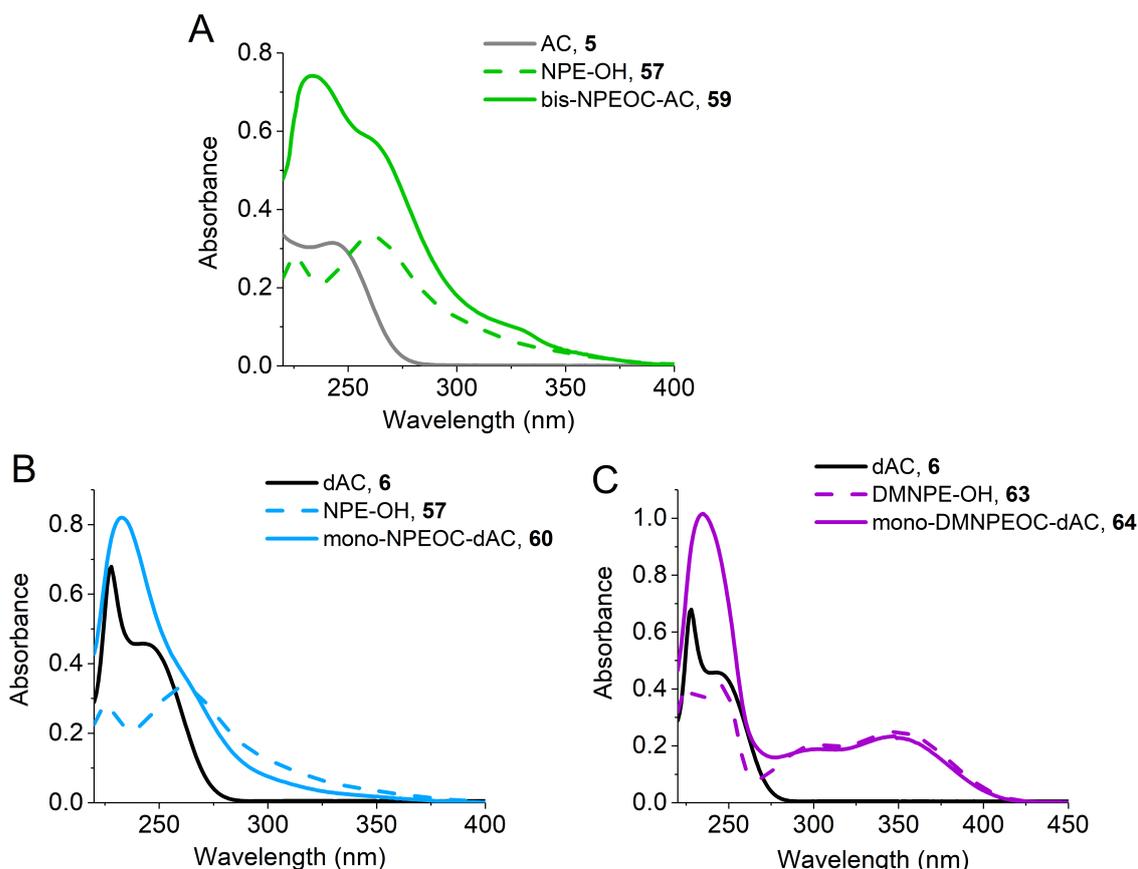


Fig. 2.1 UV-Vis absorption profiles of first generation photocaged azanucleosides

In general, all three analogues had strong absorptions in the lower UV regions. The molar absorbance of bis-NPEOC-AC **59** was $\epsilon = 14800 \text{ M}^{-1}\text{cm}^{-1}$ at $\lambda = 260 \text{ nm}$ and that of mono-NPEOC-dAC **66** was $\epsilon = 16400 \text{ M}^{-1}\text{cm}^{-1}$ at $\lambda = 233 \text{ nm}$. The strong absorption in the lower UV region is due to the triazine nucleobase. The NPE phototag absorbs modestly at $\lambda = 260 \text{ nm}$ (Figure 2.1B). However in bis-NPEOC-AC **59**, the presence of two NPE phototags significantly intensified absorption at $\lambda = 260 \text{ nm}$, represented by the additional "shoulder" in the absorption profile of **59** (Figure 2.1A). Figure 2.1C shows that the presence of two methoxy groups in DMNPE phototags

and mono-DMNPEOC-dAC **67** shifted the λ_{max} to longer wavelengths ($\lambda=365$ nm) as expected, however the absorption remained modest ($\epsilon = 4000$ M⁻¹cm⁻¹).

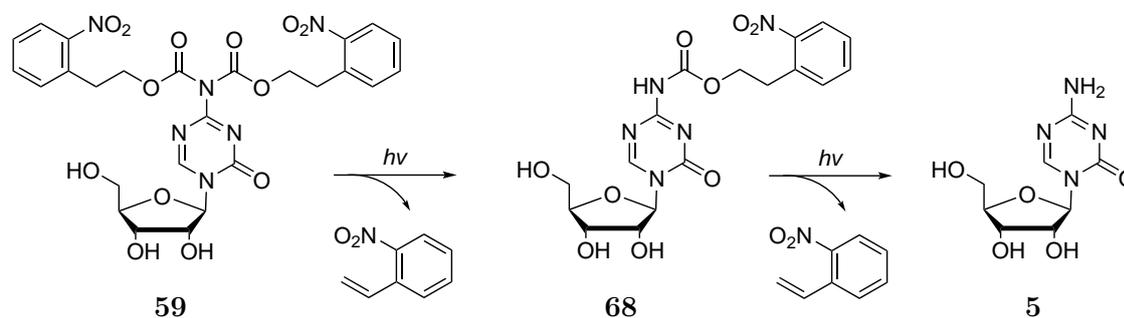
Photocaged analogue	λ_{max}	Absorption	ϵ , [M ⁻¹ .cm ⁻¹]
bis-NPEOC-AC, 59	260	0.74	14800
mono-NPEOC-dAC, 60	233	0.82	16400
mono-DMNPEOC-dAC, 64	365	0.20	4000

Table 2.1 UV-Vis absorption properties of first generation photocaged azanucleosides

Photouncaging properties

Subsequently, the photouncaging properties of the caged analogues in water were assessed. Bis-NPEOC-AC **59** was subjected to irradiation at $\lambda = 254$ nm using a benchtop UV lamp, given its strong UV absorption in this region. The uncaging reaction was monitored *via* analytical HPLC every 30 minutes over 2 hours. Analysis of the HPLC traces was challenging due to the extreme instability of triazine nucleobase in water. The caged nucleoside peak rapidly disappeared after 30 minutes, mainly giving hydrolytic and other unknown breakdown products. The addition of 1% acetic acid significantly slowed down triazine hydrolysis and enabled the analysis of two "uncaged" product peaks that were appearing over time.

The proposed uncaging mechanism is depicted in Scheme 2.6. Irradiation of **59** first gives mono-NPE-tagged **68**, which is then uncaged into AC **5**. In each step, nitrostyrene byproduct is released.



Scheme 2.6 Uncaging mechanism of bis-NPEOC-AC **59**

In Figure 2.2, the molar ratio of each species was plotted against time. After 2 hours, 80% of bis-NPEOC-AC disappeared as it deprotected into the mono-caged analogue or hydrolysed *via* the opening of the triazine ring. The amount of mono-caged analogue initially increased in response to bis-NPEOC-AC uncaging, but after 1 hour it remained unchanged.

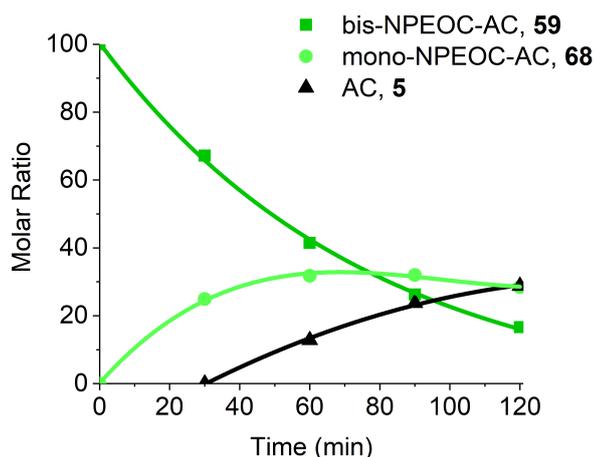


Fig. 2.2 Photouncaging profile of bis-NPEOC-AC **59** at $\lambda = 254$ nm

Although it was successfully shown that under certain conditions photoactivation occurred, the photocaged AC had several limitations. First of all, AC analogues are not viable options for uncaging as the triazine rapidly decomposed before sufficient amounts of compound could be uncaged. Furthermore, the presence of two phototags introduced an additional intermediate (mono-NPEOC-AC) that required additional uncaging into AC. This further elongated the uncaging process.

These limitations were overcome by the mono-tagged dAC analogues as the presence of a single phototag enabled direct uncaging. dAC has significantly improved aqueous stability compared to AC. The uncaging rates of mono-NPEOC-dAC **60** and mono-DMNPEOC-dAC **64** at $\lambda = 254$ and $\lambda = 365$ nm were investigated next. The results are summarised in Figure 2.3.

Monitoring the photolysis of caged dAC analogues was significantly simpler compared to caged AC analogues as triazine hydrolysis was less prevalent. At $\lambda = 254$ nm, mono-NPEOC-dAC **60** showed a much faster uncaging profile than its DMNPE analogue **64** (Figures 2.3A and C). At the end of the 90 mins irradiation, 48% of uncaged dAC was released from **60**, while **64** only released 4%. Of the initial caged compound, 18% remained in solution for analogue **60** and 58% for analogue **64**. It is likely that the

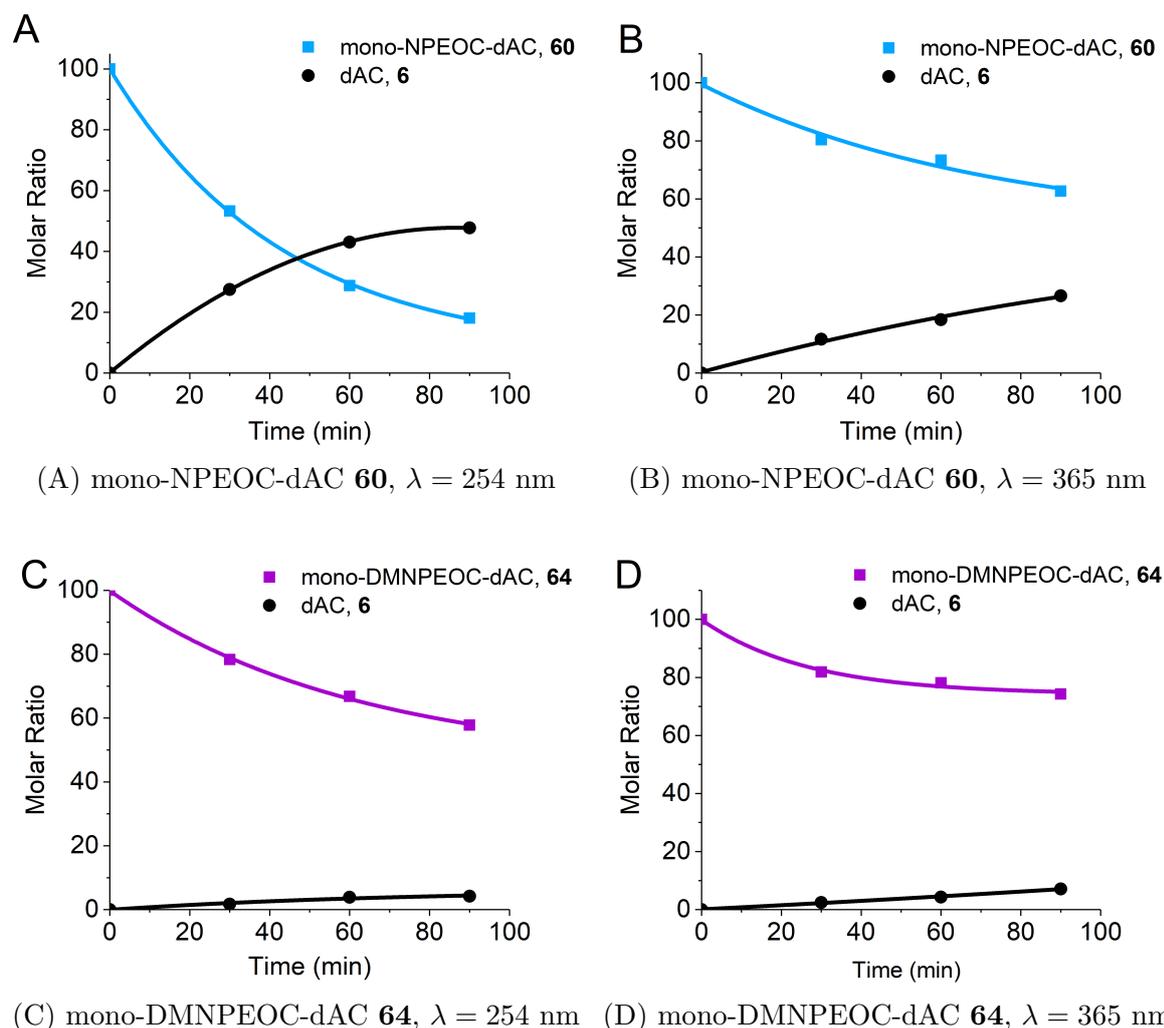


Fig. 2.3 Photouncaging profiles of first generation photocaged dAC analogues **60** and **64**

conversions into dAC at this wavelength were not quantitative, possibly due to minor hydrolysis or decomposition effects, which could be characteristics of this wavelength.

On the contrary, irradiation at $\lambda = 365$ nm avoided the decomposition effects. Conversion into dAC also seemed to be more quantitative, though photolysis occurred at a significantly slower rate. After 90 mins of irradiation, 37% of dAC was released from mono-NPEOC-dAC **60** and only 7% dAC from mono-DMNPEOC-dAC. In summary, DMNPE-caged dAC had significantly poorer uncaging rates at all wavelengths tested. Table 2.2 summarises the measured and calculated photouncaging properties of the photocaged analogues. Details of the calculations can be found in the Experimental Section (page 153).

Photocaged analogue	ϵ_{365nm} , [M ⁻¹ .cm ⁻¹]	k, [s ⁻¹]	Φ_{365nm}	$\Phi\epsilon_{365nm}$
bis-NPEOC-AC, 59	460	N/A	N/A	N/A
mono-NPEOC-dAC, 60	200	8,33E-05	4.93E-03	0.986
mono-DMNPEOC-dAC, 64	4000	6.67E-05	3.94E-03	15.76

Table 2.2 Photochemical properties of first generation photocaged azanucleosides

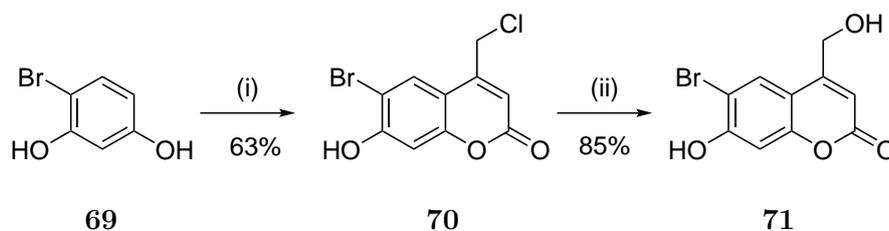
2.3 Second generation photocaged azanucleosides

While the NPE-derived phototags provided a good starting point for the caging strategy, the photochemical measurements confirmed their modest uncaging rates at $\lambda = 365$ nm. Coumarin phototags could potentially address these limitations as they have strong absorption between $\lambda = 350$ and 400 nm and reportedly higher quantum yields.

2.3.1 Design and synthesis

Two coumarin phototags were chosen for our next photocaging strategy. The 6-bromo-7-hydroxymethylcoumarin (BHC) phototag has excellent water solubility and a strong absorption at $\lambda = 370$ nm. The 7-diethylaminomethylcoumarin (DEACM) phototag has even higher λ_{max} (387 nm) and reportedly the highest quantum yields among all of the coumarin analogues,¹⁸¹ though not higher than BHC phototags' according to others.¹⁹¹

We decided to synthesise and assess both phototags before coupling them to dAC. The synthesis of the BHC phototag was carried out in two steps following literature procedures (Scheme 2.7).²¹⁵

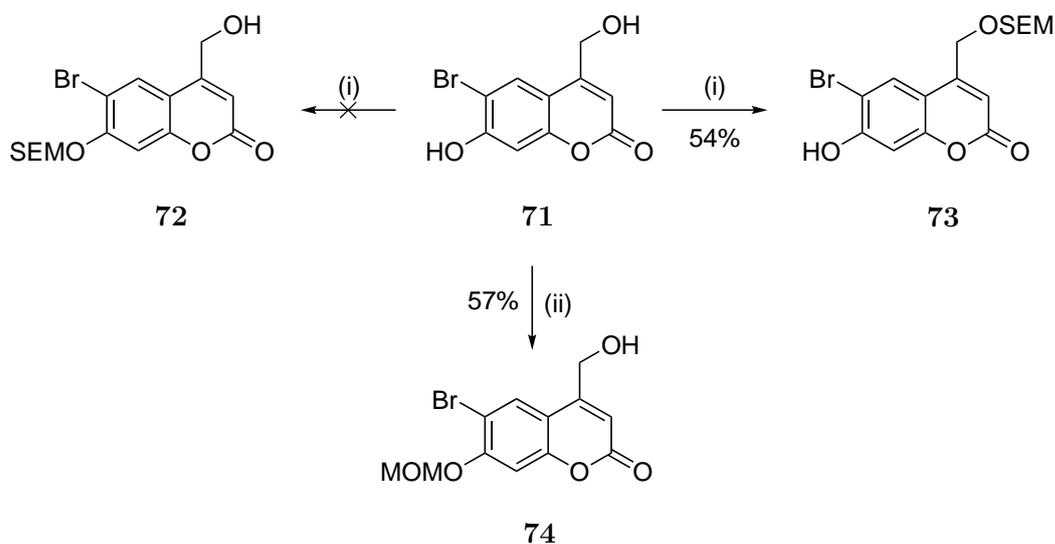


Scheme 2.7 Synthesis of BHC phototag **71**. *Reagents and Conditions:* (i) methyl 4-chloroacetoacetate, CH₃SO₃H, rt, 2 h; (ii) H₂O, reflux, 5 d

Commercially available 4-bromoresorcinol **69** was treated with methyl 4-chloroacetoacetate in the presence of methanesulfonic acid to generate 4-chloro-BHC **70** in 63% yield *via* Pechmann condensation. The resulting coumarin **70** was refluxed in water over 5 days to afford hydroxylated BHC **71** in 85% yield.

The resulting BHC phototag contained two $-OH$ groups. The aim was to convert the aliphatic $-OH$ to chloroformate to serve as a point of conjugation to azanucleosides. Protection of the aromatic $-OH$ prior to chloroformate formation was necessary to avoid any side reactions. The methoxymethyl (MOM) protecting group has been previously used to protect BHC phototags.²¹⁵ However, its harsh removal conditions using trifluoroacetic acid (TFA) can potentially interfere with the stability of azanucleosides.

As an alternative option, the 2-(trimethylsilyl)ethoxymethyl (SEM) group was used to protect the aromatic $-OH$. SEM protecting groups are structurally similar to MOM protecting groups, but their removal requires fluoride ions - a condition we know to be compatible with azanucleoside stability. BHC coumarin **71** was mixed with one equivalent of 2-(trimethylsilyl)ethoxymethyl chloride (SEM-Cl) in the presence of *N,N*-diisopropylethylamine (DIPEA) base.



Scheme 2.8 Hydroxyl-protection of BHC phototag **71**. *Reagents and Conditions:* (i) SEM-Cl, DIPEA, CH_2Cl_2 , $0^\circ C$; (ii) MOM-Cl, DIPEA, CH_2Cl_2 , $0^\circ C$

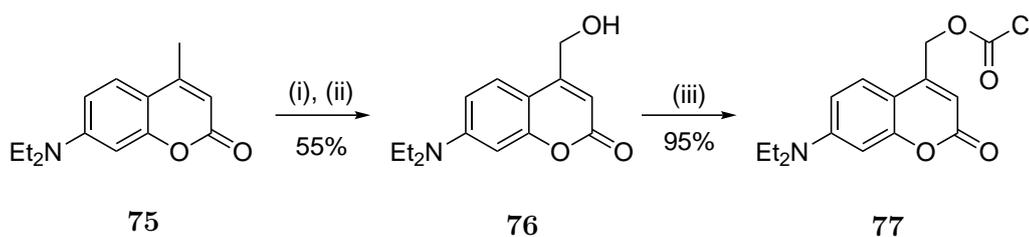
Instead of the expected analogue **72** carrying a SEM protecting group at the aromatic $-OH$, we obtained the aliphatic $-OH$ -protected version **73** in 54% yield (Scheme 2.8). Unfortunately this derivative could not be used for further coupling. MOM-protection using methoxymethyl chloride (MOM-Cl) on the other hand gave the desired MOM-

protected analogue at the aromatic –OH **74** in 57% yield (Scheme 2.8). The selectivity of the MOM-protecting groups towards aromatic and secondary alcohols could have been controlled by steric factors.

The use of BHC as a phototag for azanucleosides faced several limitations. Firstly, the BHC phototag requires the MOM protection to sufficiently couple to azanucleosides without unwanted side reactions. The removal of the MOM-protecting group after the coupling step using highly acidic conditions (as suggested in literature) may be problematic.²¹⁷ Secondly, the synthesis of phototag-chloroformate requires 4 steps including a 5-day long reaction. Furthermore the overall yield of BHC preparation was very low, suggesting that a lot of starting material would be required at the start of the synthesis to prepare sufficient amounts of phototag-chloroformate for coupling to azanucleosides.

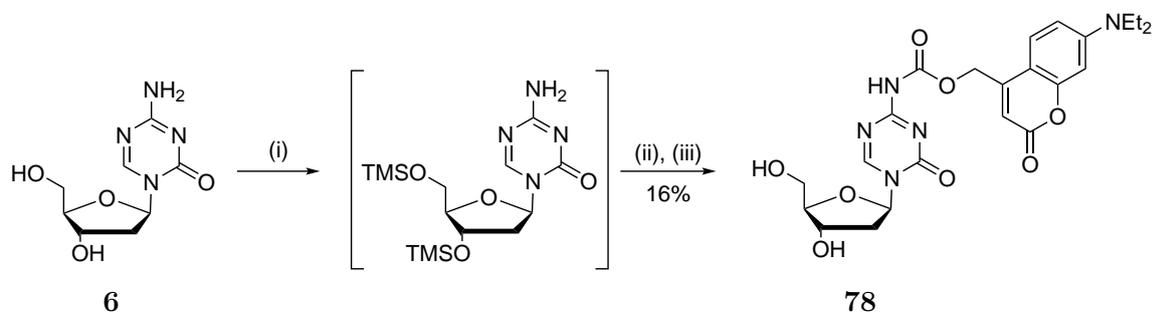
By contrast, the preparation of the DEACM phototag-chloroformate only requires 3 steps with a total duration of 4 days. Furthermore, the phototag requires no protection and the reported synthetic yields are also higher. The synthesis of DEACM phototag is summarised in Scheme 2.9.^{218,219}

Commercially available 4-diethylaminocoumarin **75** was converted into hydroxylated DEACM-OH **76** in two consecutive steps. Compound **75** was first refluxed with selenium dioxide (SeO₂) over 24 hours, followed by a 4 hour reduction with sodium borohydride (NaBH₄) to obtain **76** in 53% yield. This analogue was converted into the corresponding chloroformate **77** *via* treatment with phosgene in the presence of DIPEA over 3 hours.



Scheme 2.9 Synthesis of DEACM-chloroformate **77**. *Reagents and Conditions:* (i) SeO₂, *p*-xylene, reflux, 24 h; (ii) NaBH₄, EtOH/THF, rt, 4 h; (iii) COCl₂, DIPEA, THF, 0 °C, 3 h

The resulting DEACM-chloroformate **77** was subsequently coupled to dAC **6** *via* the transient protection strategy to obtain N-DEACMOC-dAC **78** in 16% yield (Scheme 2.10).



Scheme 2.10 Synthesis of N-DEACMOC-dAC **78**. *Reagents and Conditions:* (i) HMDS, pyridine, DMF rt, 2h; (ii) 2 eq chloroformate **77**, CH₂Cl₂/DMF, rt, 16 h; (iii) TAS-F, DMF, rt, 2 h

2.3.2 Photochemical characterisation

N-DEACMOC-dAC **78** exhibited strong absorption between $\lambda = 350$ and 450 nm, with $\lambda_{max} = 387$ nm (Figure 2.4A). Its molar absorbance at $\lambda = 395$ nm was $\epsilon = 10000$ M⁻¹cm⁻¹ and at $\lambda = 365$ nm it is $\epsilon = 7000$ M⁻¹cm⁻¹, almost twice the absorbance of mono-DMNPEOC-dAC **64** ($\epsilon = 4000$ M⁻¹cm⁻¹). Figure 2.4B shows the photouncaging profile of **78**.

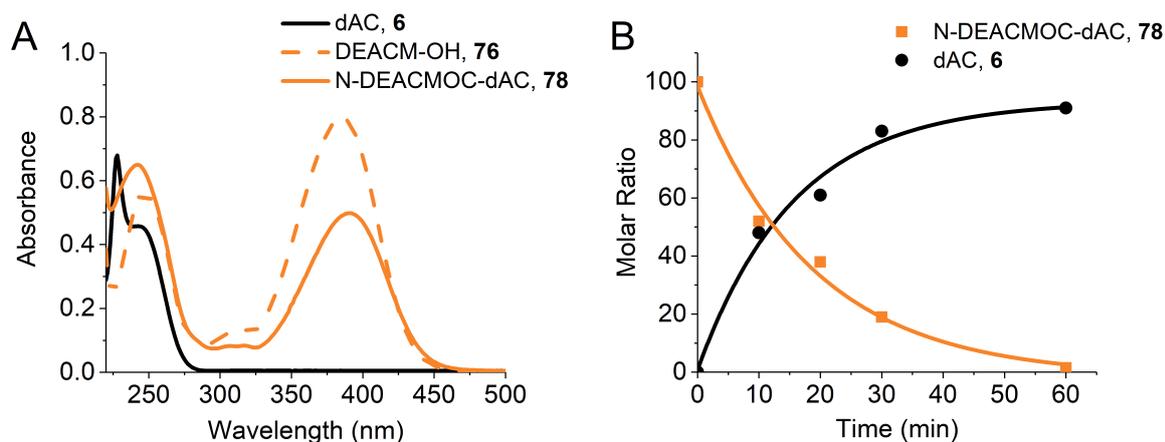


Fig. 2.4 (A) UV-Vis absorption and (B) photouncaging profile of N-DEACMOC-dAC **78**

The strong absorption at $\lambda = 365$ nm is reflected in the rate and efficiency of the photolysis. After 1 hour of irradiation at $\lambda = 365$ nm, almost all of the caged compound (98%) has disappeared, and 'uncaged' dAC appeared in almost quantitative amounts (92%). This uncaging profile is significantly faster compared to the NPE- and DMNPE-tagged azanucleosides that were examined previously.

The HPLC traces for the photolysis of analogue **78** at $\lambda = 365$ nm are shown in Figure 2.5. The traces show that the single peak ($R_t = 14$ min) representing the caged analogue **78** at $t = 0$ min rapidly disappeared after 30 and 60 min. This was accompanied by the quantitative appearance of two additional peaks that corresponded to dAC ($R_t = 4.5$ min) and DEACM-OH ($R_t = 15$ min). The traces show that the uncaging proceeded without any unwanted side products. The calculated quantum yield of uncaging ($\Phi_{365nm} = 0.028$) was lower than reported values in literature.¹⁹¹

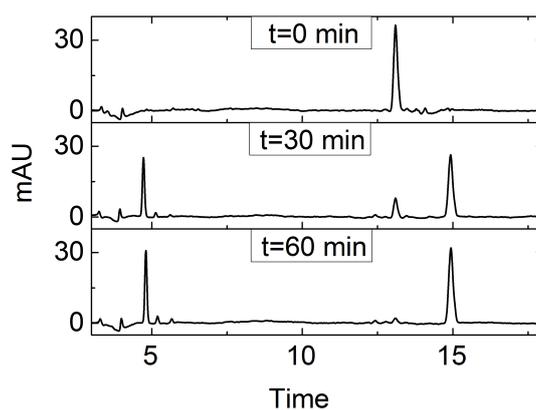


Fig. 2.5 HPLC traces for photouncaging of N-DEACMOC-dAC **78** at $\lambda = 365$ nm

2.3.3 Stability studies

In summary, N-DEACMOC-dAC **78** had the best photochemical characteristics among all the photocaged azanucleoside analogues. Since this analogue was prepared for the purpose of cell biological applications, it was essential that it possessed good stability under such conditions.

The stability of **78** was assessed by incubation in water at 37°C in the dark and aliquots were analysed *via* HPLC. The results indicated that compound **78** had very poor stability (Figure 2.6). After 2 hours of incubation, nearly all of the caged

compound **78** had decomposed followed by complete degradation after 5 hours. It was also observed that the main peak was degrading into two unknown peaks.

To confirm the identity of the peaks, the decomposed solution was spiked with dAC and re-analysed by HPLC. dAC eluted after peak 1, suggesting that peak 1 may be a more polar, closely related analogue of dAC. The solution was further spiked with DEACM-OH to determine the identity of peak 2. HPLC analysis of the re-spiked mixture showed that DEACM-OH and peak 2 identical retention times, indicating that peak 2 corresponds to DEACM-OH.

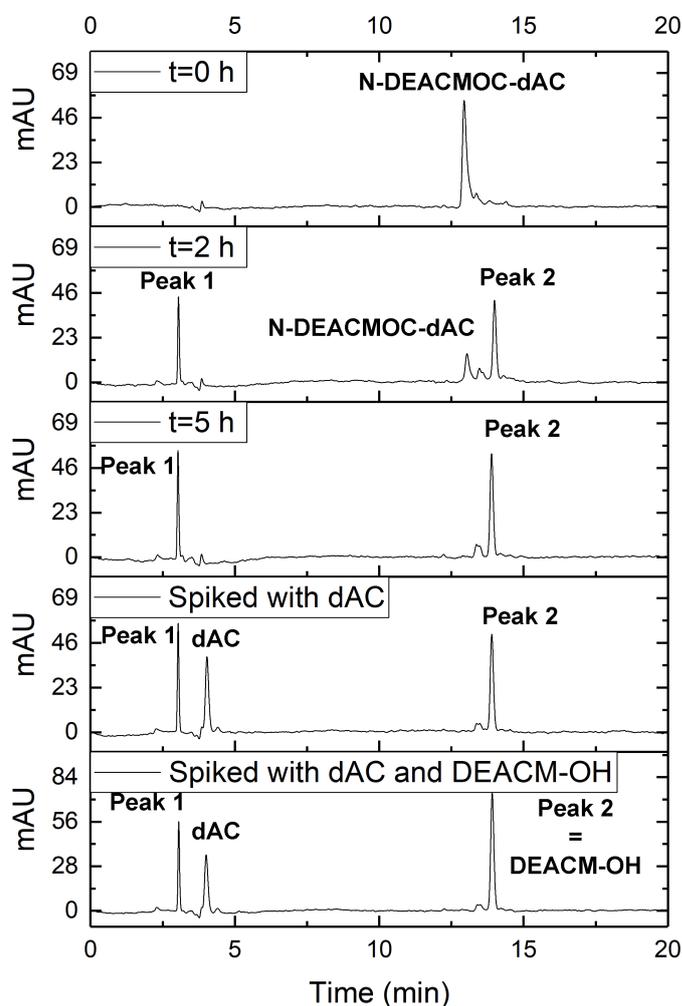
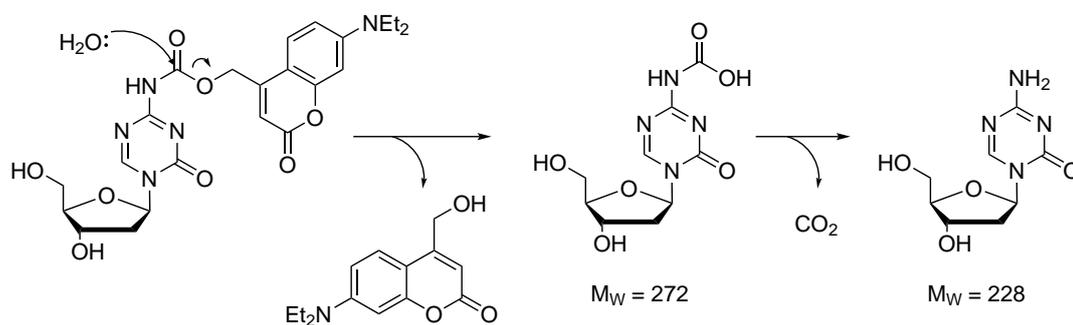


Fig. 2.6 HPLC stability studies of N-DEACMOC-dAC **78** at 37°C. In order to determine the identity of the degradation peaks, the stability mixture was spiked with dAC, followed by DEACM-OH.

Together these findings suggested that the dark instability was due to the carbamate bond hydrolysis that resulted in the dissociation of the phototag. One possible mechanism for the process is outlined in Scheme 2.11. The carbamate linker between the phototag and the azanucleoside can dissociate in the presence of water, leading to the release of the DEACM-OH phototag and potentially a carbamic acid intermediate that could eventually decarboxylate into dAC. The rate of decarboxylation is dependent on a range of factors, such as temperature and pH of the solution and the characteristics of the carboxylated compound itself.



Scheme 2.11 Proposed mechanism for dark instability of N-DEACMOC-dAC **78**

The carboxylated dAC is a viable intermediate for this breakdown process as the HPLC analysis suggests the presence of the dissociated DEACM-OH phototag (corresponding to peak 2 in Figure 2.6) and a more polar dAC analogue (corresponding to peak 1 in the same Figure). In order to test this hypothesis, further characterisation of the degradation products was attempted by mass spectrometry.

In the mass spectrometry analysis, the stability mixture was re-created (1 mM of N-DEACMOC-dAC, **78** was incubated in water overnight) and peak 1 of the stability mixture was isolated via semi-preparative HPLC. The solvent was removed and the mixture was submitted for mass spectrometry analysis. The mass spectrometry analysis was carried out by Dr Kersti Karu.

The calculated exact mass of the carbamic acid intermediate is 272.08 and of the dAC is 228.09. The negative mode electrospray ionisation analysis (TOF MS ES⁻) detected a base peak value at 226.4376 as shown in Figure 2.7, which is likely corresponding to dAC (that has lost two protons of 3' and 5'-OH during the negative mode ionisation). This suggests that the analysed mixture predominantly contained dAC.

Another significantly smaller peak at 270.4261 m/z could potentially correspond to the carbamic acid intermediate. However, it is likely that the carbamic acid intermediate

would lose 3 protons (given the low pKa of carbamic acid derivatives) during the electrospray ionisation process, giving an expected mass of 269 after ionisation, which is one less than the value detected during the analysis. Given this, the presence of the carbamic acid intermediate cannot be confirmed from these measurements.

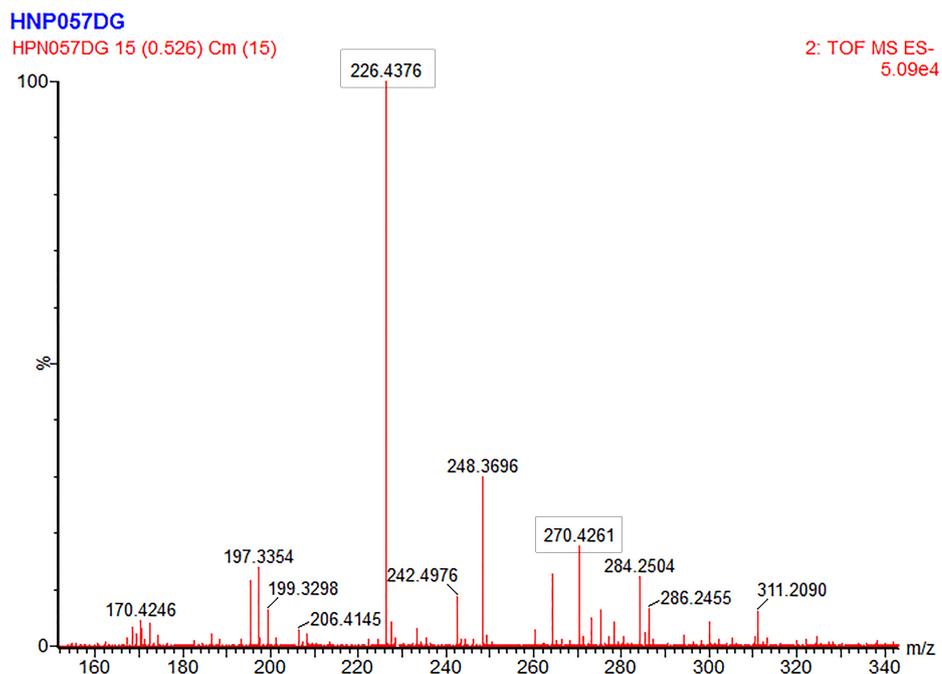


Fig. 2.7 Mass spectrometry profile of N-DEACMOC-dAC **78** degradation product

The HPLC and mass spectrometry analyses gave significantly different observations, suggesting different mechanisms of the carbamate instability. The HPLC measurements suggested that when the carbamate bond was hydrolysed by water, a more polar, (likely carbamic acid intermediate) of dAC was formed, which can potentially convert into dAC by decarboxylation. The kinetics of the latter process was not measured by HPLC. On the other hand, mass spectrometric analysis of the isolated intermediate suggested that it mainly composed of dAC and the presence of carbamic acid dAC could not be confirmed.

There could be several reasons for this discrepancy between the HPLC and Mass Spectrometry analyses. While the HPLC analysis gave a direct measurement of the stability mixture that only had to pass through a C18 column until detection of the individual compound peaks, the sample for the mass spectrometry analysis had to go several steps of preparation. Starting from the same stability mixture, the carbamic acid intermediate was first isolated by semi-preparative HPLC, followed by solvent removal

using a rotary evaporator. The dried sample was then resolubilised and subjected to mass spectrometry analysis. During the mass spectrometry analysis, the compound underwent electrospray ionisation, that typically results in the deprotonation of the sample in order to enhance their detection. Given the greater number of preparative steps for the mass spectrometry analysis, it is possible that the compound (that is already supposedly unstable, based on its structure), underwent decarboxylation during one of these processes which then resulted in the dominating presence of dAC in the final sample mixture.

Both the photouncaging and stability experiments conducted on N-DEACMOC-dAC, **79** resulted in the release of the photocage (and eventually dAC), via decarboxylation. However, HPLC analyses indicated that the mechanism of these two processes are different, given the different products observed:

During the photouncaging, light irradiation targets the C-O bond of the photocage, leading to heterolytic bond cleavage (as shown and explained in Scheme 1.12). This is followed by the formation of a tight ion pair, the positively charged photocage and the negatively charged carbamate dAC intermediate. Once water attacks the positively charged photocage, the negatively charged dAC carbamate undergoes decarboxylation to give the final dAC product that is detected by HPLC.

During the hydrolytic instability, the presence of water molecules trigger the process by first attacking the carbamate bond, which then releases the photocage (confirmed by HPLC). The remaining carbamic acid intermediate is still likely protonated and this could be a reason why decarboxylation does not happen perhaps instantly. Due to that, the carbamic acid analogue can still be detected via HPLC, even after 3 hours of incubation. On the contrary, the series of preparative steps in mass spectrometry analysis may have contributed to the deprotonation and therefore accelerated decarboxylation of the carbamic acid analogue which may have lead to the major presence of dAC in the isolated mixture.

In the future, further work is prompted to better understand the main factors (e.g. sample preparation process, stability solution environment) on the deprotonation and decarboxylation properties and kinetics of the carbamic acid intermediate to better predict the stability profile of caged-dAC compounds.

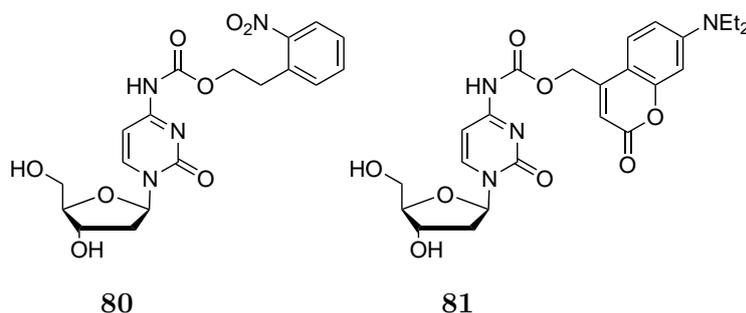
Further analysis of the dark instability of **79** revealed that the rate of degradation decreased when other solvents such as acetonitrile (ACN) were present in the stability solution (Table 2.3). Furthermore, when biological buffers were used instead of water, the rate of degradation increased.

Time (min)	0.1% ACN/Water 27 °C	0.1% ACN/ Water 37 °C	15% ACN/ Water 37 °C	MOPS-KOH pH 7.2 37 °C	HEPES pH 7.5 37 °C
0	100	100	100	100	100
30	81	62	77	52	31
60	78	45	62	28	15

Table 2.3 Stability of N-DEACMOC-dAC **78** under alternative solvents and buffers

To investigate the cause of the dark instability, the stabilities of alternative analogues were explored. The aim was to establish whether the cause of the instability was dAC, the carbamide bond or the phototag itself. To test the latter, the stability of another photocaged dAC analogue, mono-NPEOC-dAC **60** was measured. In addition, the stabilities of other nucleosides were assessed. We chose dC as our reference compound due to its closely related structure dAC. The nucleobase of dC contains a carbon instead of a nitrogen in position 5, which makes it non-susceptible to ring opening.

dC reference compound was caged with NPE and DEACM phototags to yield the two photocaged dC analogues mono-NPEOC-dC **80** and mono-DEACMOC-dC **81** (Their syntheses are described in the Experimental Section) (Scheme 2.12).



Scheme 2.12 Structures of photocaged dC reference compounds **80** and **81**

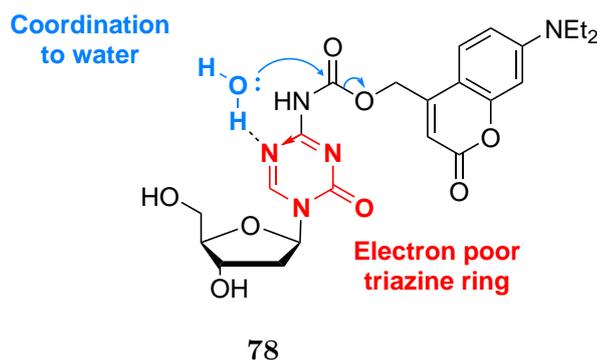
The stability studies suggested that the phototag had a moderate effect on compound stability (Table 2.4). Mono-NPEOC-dAC **60** was more stable than DEACM-caged dAC **78** (after 60 mins of incubation, 85% of **60** remained intact compared to 45% of

78). However, both caged dAC compounds showed considerable dark instability. By contrast, it was also observed that dC analogues had excellent stability at 37°C even after 7 hours of incubation.

Time (min)	% DEACMOC-dAC, 78	% DEACMOC-dC, 81	% NPEOC-dAC, 60	% NPEOC-dC, 80
0	100	100	100	100
30	62	100	91	100
60	45	100	84	100

Table 2.4 Stability results of photocaged dAC analogues and dC reference compounds

The results indicate that the carbamate bond alone is stable to hydrolysis. It was confirmed that the replacement of a C to a N atom in the triazine ring had a dramatic effect on the stability of photocaged dAC analogues. The possible explanation for this phenomenon is summarised in Scheme 2.13. The nitrogen atom in position 5 may have an inductive effect on the triazine ring. As a result, the triazine ring could become electron poor, making the neighbouring carbamate bond more susceptible to hydrolytic attack. In addition, the nitrogen atom in position 5 may coordinate to water molecules *via* its lone pairs. The presence of water molecules in such close proximity to the carbamate bond may also increase the chance of hydrolytic attack.



Scheme 2.13 Proposed mechanism of instability in N-DEACMOC-dAC **78**

In summary, we successfully photocaged azanucleosides at the nucleobase position. Unfortunately the incubation of these analogues at 37 °C led to rapid dissociation of the phototag due to the electronic properties of the triazine ring. We proposed that placing the phototag on alternative positions would overcome these limitations.

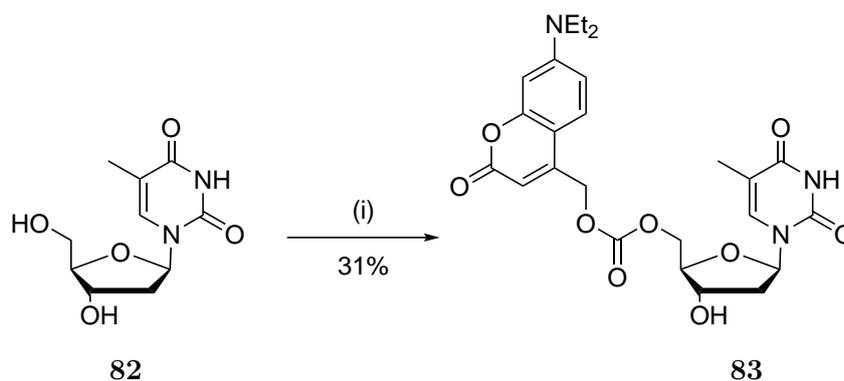
2.4 Third generation photocaged azanucleosides

Given the the strong influence of the triazine ring on compound stability, we envisaged that placing the photocage away from the triazine ring could improve compound stability. Therefore the aim was to place the phototag on the 5'-OH and 3'-OH of dAC to avoid the direct link between the phototag and the triazine ring. In this section, we will describe the preparation and characterisation of 5'- and 3'-caged dAC analogues.

2.4.1 Pre-assessment of carbonate linker stability

The proposed strategy had the consequence of replacing the carbamate linker with a carbonate linker as the photocage would be linked to an -OH instead of a -NH₂ group. Carbonate linkers have reportedly poorer hydrolytic stabilities,²²⁰ though some studies claimed otherwise in the context of 5'-caged nucleosides.²²¹ With regard to DEACM phototags, no stability data of 5'-DEACM-caged nucleosides have been reported to date. The aim was to assess the stabilities of other 5'-DEACM-caged nucleosides (attached *via* a carbonate linker) to determine if these linkers have stability issues.

DEACM phototag was first coupled to thymidine nucleoside to assess the carbonate link stability. Thymidine is a good starting point as it does not require any nucleobase protection prior to coupling. 5'-caging of thymidine was achieved in one step, *via* direct coupling of DEACM chloroformate **77** to thymidine **82** following a protocol described by Bühler *et al.* (Scheme 2.14).²⁰³ The resulting 5'-DEACMOC-thymidine **83** was isolated in 31% yield. Stability analysis of the compound indicated that it is stable even after 7 hours of incubation at 37 °C (98% of the caged-compound remained intact). This finding confirms that the carbonate linker and the DEACM-phototag do not cause inherent stability issues.

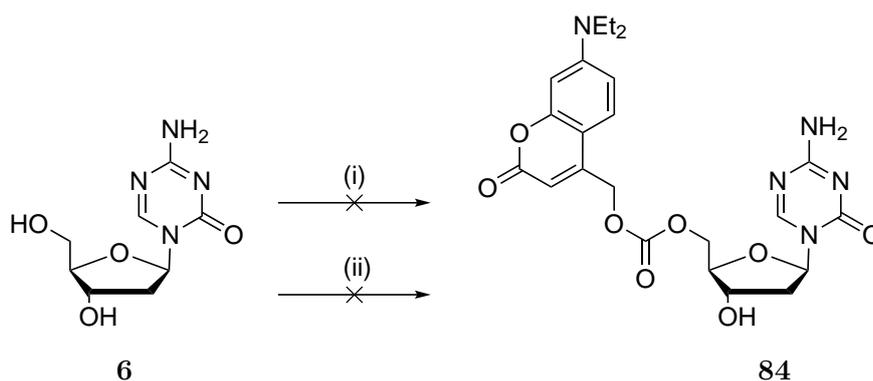


Scheme 2.14 Synthesis of 5'-DEACMOC-Thymidine **83**. *Reagents and Conditions:* (i) 1.3 eq chloroformate **77**, CH₂Cl₂/pyridine, rt, 16 h

2.4.2 Synthesis of 5'-caged analogue

Direct coupling of DEACM-chloroformate to dAC

For the synthesis of 5'-caged dAC, direct coupling of dAC to the DEACM-phototag was first attempted. Due to the low reactivity of the exocyclic amine, it was assumed that its separate protection was not necessary, given that the 5'-OH is the most reactive among the three functional groups of dAC. Unfortunately direct coupling of DEACM-chloroformate **77** to unprotected dAC **6** did not yield any product even in the presence of excess (5 equivalents) phototag.



Scheme 2.15 *Reagents and Conditions:* (i) 5 eq chloroformate **77**, CH₂Cl₂/pyridine, rt, 16 h, (ii) 1.5 eq chloroformate **77**, DBU, DMF, rt, 16 h

An alternative strategy was pursued. A recent study by Weiss *et al.* described the 5'-functionalisation of gemcitabine without the protection of the exocyclic amine.²²² Gemcitabine is a closely related structural analogue of dC with a fluorine group attached

to the C2 position of the ribose ring. Direct coupling to the 5'-OH of gemcitabine was achieved in the presence of a bulky base 1,8-diazabicycloundec-7-ene (DBU). The method was trialled on DEACM-chloroformate **77** and dAC but unfortunately no product was detected. Due to the unsuccessful attempts to directly couple photocage chloroformate to dAC, the protection of the exocyclic amine was explored next.

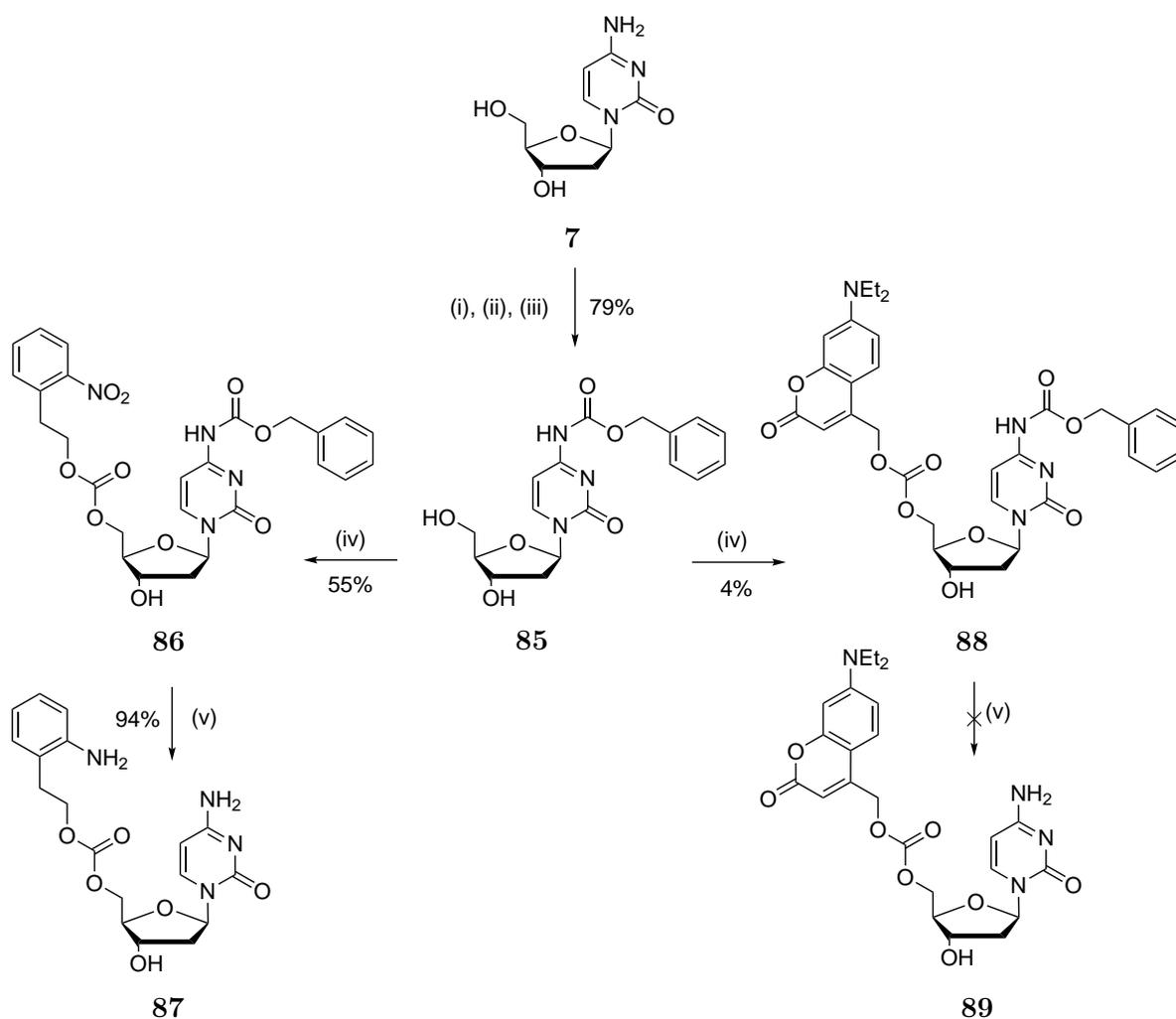
5'-coupling *via* protection of the exocyclic amine

In order to develop an optimised protocol for the synthesis, commercially available dC **7** was used as a model compound. The synthetic strategy is outlined in Scheme 2.16. The first step of the synthesis was the protection of the amine nucleobase with carboxybenzyl (CBz) protecting group. This was achieved in three steps, following the transient protection strategy to obtain N-Cbz-dC **85** in 79% yield. The CBz group was chosen as the protecting group because its removal conditions are mild enough to leave dAC and the triazine ring intact. Although the optimisation of the synthetic protocol was carried out using dC **7**, all of the reagents and reaction conditions were chosen with careful consideration to ensure their compatibility with dAC stability.

The next step was to conjugate a phototag on the 5'-position. Analogue **85** was first coupled to NPE-chloroformate **58** to obtain N-Cbz-5'-NPEOC-dC **86** in 55% yield. In the final step, the CBz protecting group was rapidly removed *via* palladium-assisted hydrogenation to give 5'-caged compound in 94% yield (although the nitro group of the phototag was also reduced to an amine group, in line with expectations). This confirmed that the strategy is feasible, so the coupling of the DEACM-phototag was trialled next.

Coupling of Cbz-dC **85** to DEACM-chloroformate **77** to obtain N-Cbz-5'-DEACMOC-dAC **88** was successful, but with significantly lower yield (4%). The final hydrogenation of analogue **88** was unsuccessful, due to the nitrogen lone pair of $-\text{NEt}_2$ that may have coordinated to and interfered with the palladium catalyst.

In summary, the synthesis of 5'-caged analogue was not possible *via* CBz protection of the nucleobase. The removal of the CBz group was deemed challenging in the presence of DEACM-phototag. Alternative nucleobase protecting groups in literature (benzoyl, isobutyryl and dimethylformamidyl) are generally removed under very harsh conditions (conc. NH_4OH , 55°C). And even the recent, ultramild protecting groups (phenoxyacetyl-, acetyl- groups) are removed under basic conditions ($\text{MeOH}/\text{K}_2\text{CO}_3$), which would be detrimental to dAC stability. Given the absence of any alternative,



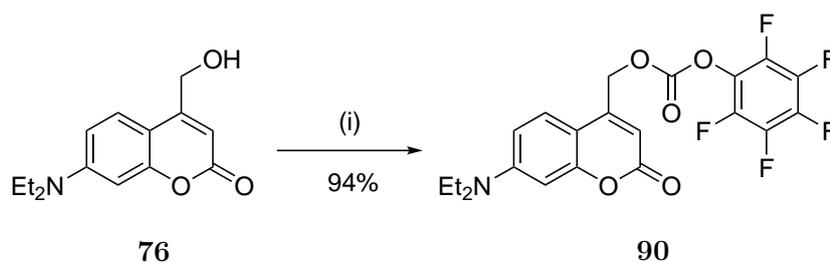
Scheme 2.16 Synthesis of 5'-caged dC *via* protection of exocyclic amine. *Reagents and Conditions:* (i) HMDS, pyridine/DMF, rt, 2 h; (ii) Cbz-Cl, CH₂Cl₂/pyridine, rt, 16 h; (iii) TAS-F, DMF, rt, 2 h; (iv) NPE-chloroformate **58**, CH₂Cl₂/pyridine, rt, 16 h; (v) H₂/Pd, MeOH; (vi) DEACM-chloroformate **77**, CH₂Cl₂/pyridine, rt, 16 h

viable protecting groups that could be tested, an alternative strategy was pursued to synthesise 5'-caged dAC analogues.

PFP-activated and DMAP-catalysed direct coupling

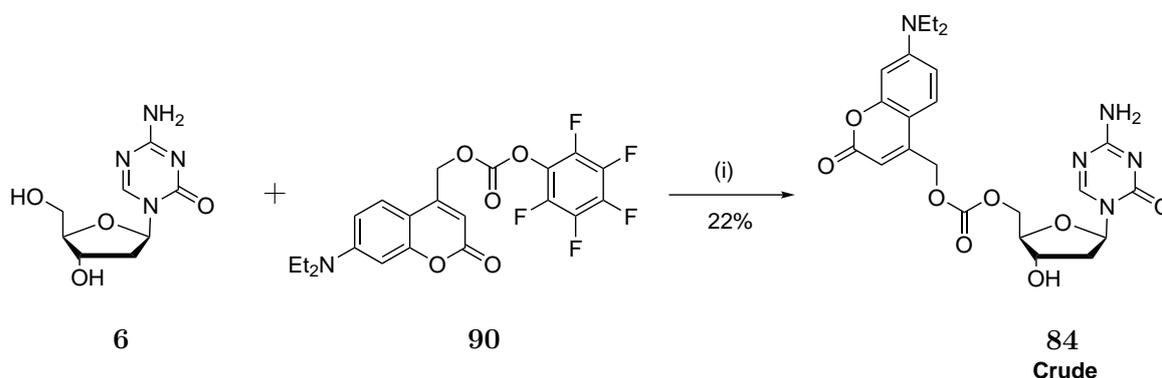
The next strategy relied on the coupling of DEACM phototag to 5'-OH of dAC without any protecting groups. Since the chloroformate on its own had limited reactivity towards 5'-OH, an alternative activation and coupling strategy was required. In a recent study by Zhang *et al.*, the synthesis of 5'-dAC esters were described.¹⁷³ The authors coupled activated amino acid moieties to dAC in the presence of 4-(dimethylamino)pyridine (DMAP) catalyst to form the desired dAC ester. It was envisaged that DEACM-phototag could be similarly activated and coupled to dAC *via* a carbonate bond instead of an ester bond.

The pentafluorophenyl (PFP) group was chosen as the activating group for DEACM phototag. PFP groups have been previously used to activate and couple amino acids in peptide synthesis, but they can also be used to activate alcohols. PFP-esters are sufficiently stable to isolate and they generally undergo couplings with high efficiencies.^{223,224} DEACM-OH alcohol was treated with bis(pentafluorophenyl)carbonate (BPFPC) in the presence of DIPEA to obtain PFP-activated DEACM in 94% yield (Scheme 2.17).



Scheme 2.17 Synthesis of DEACMOC-PFP **90**. *Reagents and Conditions:* (i) BPFPC, DIPEA, THF, rt, 16 h

The PFP-activated DEACM-carbonate **90** was coupled to dAC **6** in the presence of DMAP coupling agent (Scheme 2.18). After overnight stirring, the formation of a new compound was detected by thin layer chromatography (TLC) and was successfully isolated at 27% as a yellow solid. Mass spectrometry analysis indicated that this analogue had the desired mass, identical to the target compound 5'-DEACMOC-dAC **84**.



Scheme 2.18 Synthesis of 5'-DEACMOC-dAC **84** via DMAP coupling and PFP activation. *Reagents and Conditions:* (i) DMAP, DMF, rt, 16 h

NMR analysis in DMSO confirmed the presence of DEACM peaks in addition to the dAC peaks. It was also observed that the 5'-OH peak that is normally present in dAC had disappeared, indicating that the 5'-OH group has been successfully modified. However, there were additional peaks present that were yet to be identified. HPLC analysis of the solid revealed the presence of two main peaks in the mixture despite appearing as a single spot on TLC (Figure 2.8). The different batches correspond to different fractions purified by flash chromatography.

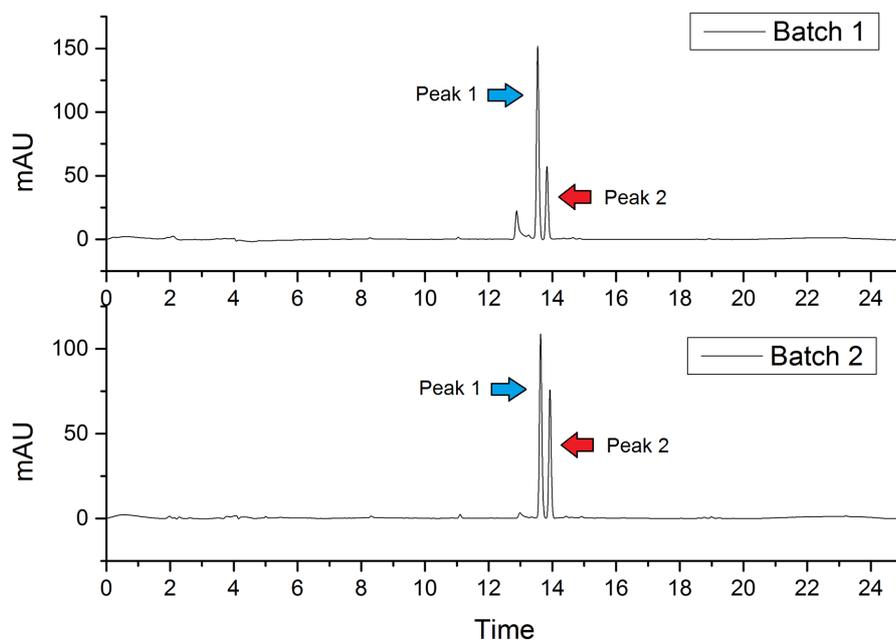
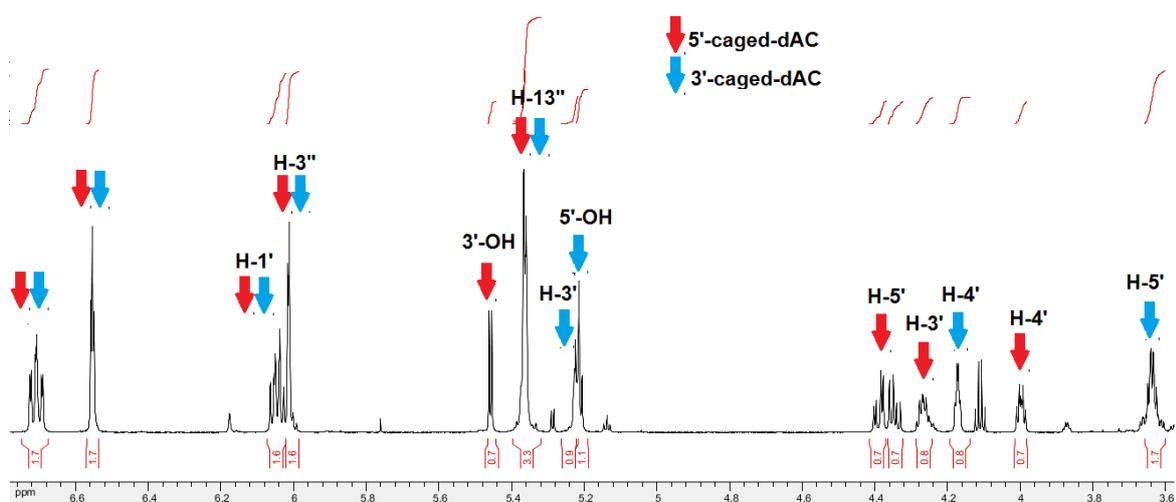
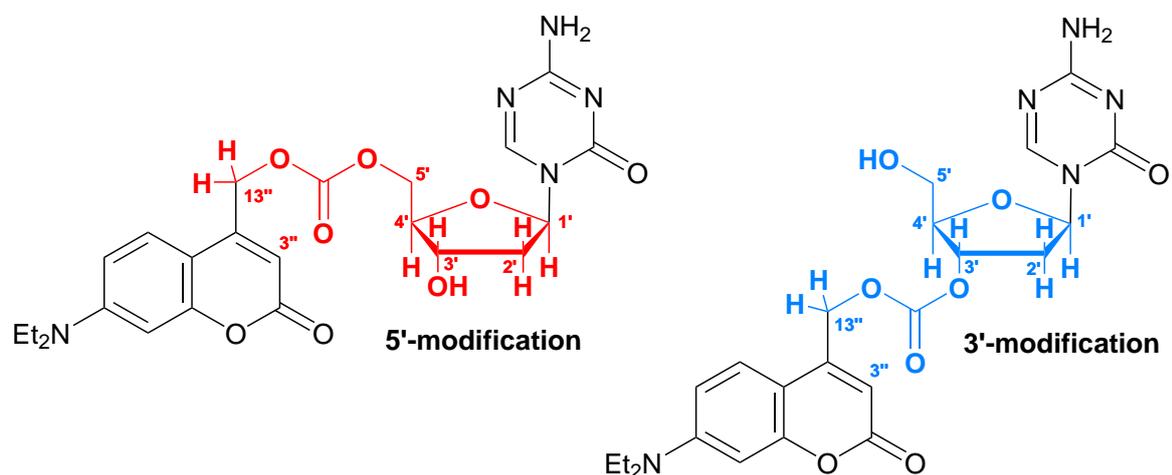


Fig. 2.8 HPLC traces of purified 5'-DEACMOC-dAC 'crude'

Further NMR analysis of the mixture indicated that the additional peaks belonged to the 3'-modified dAC analogue (Figure 2.9A). Most notably, the 3'-OH peak was absent and the proton peaks from the deoxyribose ring had shifted δ values: the value for H-3' was higher and H-5' was lower. Other peaks (H-1', triazine and DEACM peaks) remained unchanged. This suggests that during the coupling, modification at the 5'-OH and 3'-OH occurred simultaneously despite the lower reactivity of 3'-OH.



(A) Partial NMR spectrum of the crude mixture with annotated H environments



(B) Structures of 5'- and 3'-modified dAC analogues with highlighted H environments corresponding to annotated NMR peaks in (A)

Fig. 2.9 NMR analysis of 'crude' mixture from the synthesis of 5'-DEACMOC dAC 84

Optimisation of synthesis and purification of 5'-caged analogue

Isolation of pure 5'-caged analogue from the 5'- and 3'-caged dAC mixture was next attempted. The first issue was that the two compounds had identical R_f values on TLC and were eluted as one single spot. However, it was possible to analytically separate 3'- and 5'-caged dAC *via* HPLC, though the two peaks eluted soon after each other ($R_f = 13.6$ and 13.8 min). HPLC was essential to differentiate between the two analogues. Therefore the purification of these mixtures involved flash column chromatography coupled to HPLC analysis to assess the purity of the collected fractions.

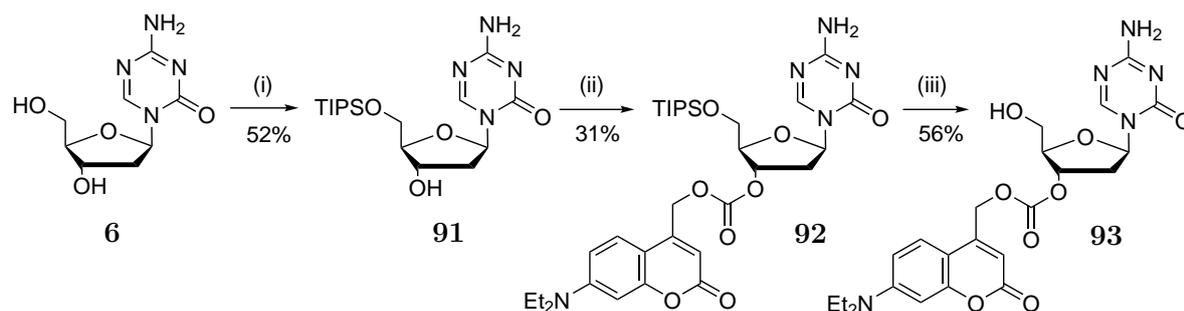
In order to optimise the synthesis of 5'-caged analogue, the reaction conditions were modified to rely on the higher reactivity of 5'-OH compared to 3'-OH. The synthesis was repeated at $-20\text{ }^\circ\text{C}$ to slow down the reaction and enhance the equilibrium towards the formation of the kinetic product, the 5'-caged analogue. In addition, only 0.8 equivalent of the phototag was used and the reaction time was reduced to 5 h. When isolating the products of these reactions, it was found that the ratio between 5'- and 3'-caged analogues was improved (5:1). Careful column purification with slow MeOH gradient allowed the collection of 5 mg (7% yield) of pure product which was only enough for NMR and MS analyses.

Since selective synthesis of 5'-caged dAC was unsuccessful and there were no alternative 3'-OH specific protecting groups, the next aim was to optimise the purification of the 5'- and 3'-caged dAC mixture. Conventional flash column purification did not yield sufficient amounts of pure 5'-caged analogue, semi-preparative HPLC was explored as a purification method. Yet it did not yield the desired outcome as peak separation was limited. The amount that could be purified from this method was also low.

Recrystallisation was the final purification method that was attempted. The synthesis was repeated at a larger scale to obtain sufficient amounts of crude mixture for recrystallisation trials 5'-DEACMOC-dAC **84** was successfully recrystallised from methanol, collected in sufficient amounts (>25 mg) and had its purity confirmed by HPLC.

2.4.3 Synthesis of 3'-caged analogue

Following the conjugation of DEACM phototag to the 5'-position, photocaging strategies at the 3'-position were also explored. The synthesis of the 3'-caged analogue was achieved in three steps (Scheme 2.19).



Scheme 2.19 Synthesis of 3'-DEACMOC-dAC **93**. *Reagents and Conditions:* (i) TIPS-Cl, imidazole, DMF, rt, 16 h; (ii) DEACMOC-PFP **90**, DMF, 0 °C; (iii) AcCl, MeOH, rt, 3 h

In order to selectively synthesise 3'-modified analogue, the 5'-OH was first protected with the triisopropylsilyl (TIPS) protecting group. TIPS is one of the most stable silyl protecting groups in literature. Apart from its great stability, it was chosen due to its mild, fluoride-based removal protocol that has previously shown compatibility with dAC stability.

dAC **6** was treated with triisopropylsilyl chloride (TIPS-Cl) in the presence of imidazole base to afford 5'-TIPS-dAC **91** in 52% yield. The TIPS-protected dAC was treated with PFP-activated DEACM phototag **90** in the presence of DMAP catalyst. These are the same conditions used for the 5'-caged analogue.

The final step was deemed to be more challenging than initially expected. Several unsuccessful attempts were made to remove the TIPS-protecting group using TAS-F and TBAF. An alternative protocol reported by Khan *et al.* was utilized.²²⁵ Analogue **92** was treated with acetyl chloride in anhydrous methanol (to generate anhydrous hydrogen chloride *in situ*) in the presence of molecular sieves for 3 hours. The desired product 3'-DEACMOC-dAC **93** was successfully isolated in 56% yield.

2.4.4 Photochemical characterisation

UV-Vis absorption

The UV-Vis absorption properties of the 3'- and 5'-caged analogues were comparable to that of the N-caged mono-DEACMOC-dAC's (summarised in Figure 2.10).

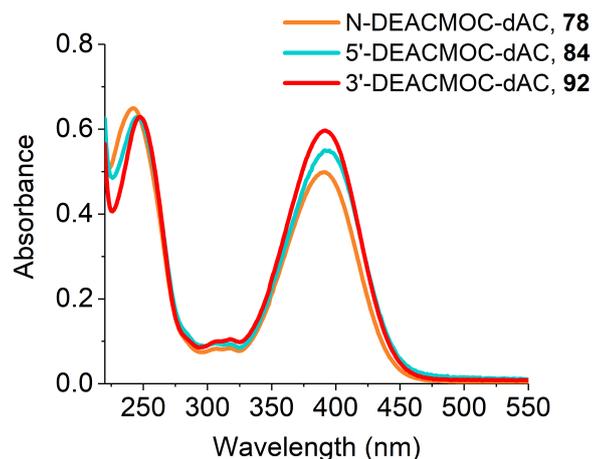


Fig. 2.10 UV-Vis absorption profiles of DEACMOC-caged analogues **78**, **84** and **93**

Absorption at $\lambda = 250$ nm for all three compounds were almost identical, with N-caged analogue having slightly higher absorption compared to the other two. All compounds had strong absorption around $\lambda = 400$ nm, but the extinction coefficients varied between the analogues. N-caged analogue had the lowest absorption ($\epsilon = 10000 \text{ M}^{-1}.\text{cm}^{-1}$), whereas 3'-DEACMOC-dAC had the highest ($\epsilon = 12000 \text{ M}^{-1}.\text{cm}^{-1}$).

The slight variation in absorption could be accounted for by the positioning of the phototag on the dAC nucleoside. The absorption coefficients of the caged analogues increased as the phototag was placed further from the triazine ring. It is possible that the triazine ring affects the absorption of the DEACM-ring *via* electronic effects.

Photouncaging properties

The uncaging profiles of the 5'- and 3'-caged analogues at $\lambda = 365$ nm are summarised in Figure 2.12.

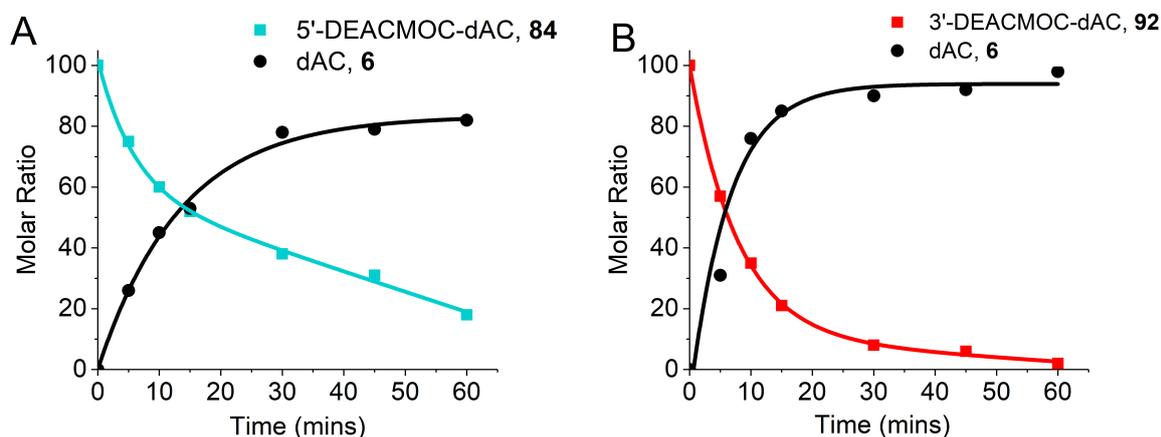


Fig. 2.11 Photouncaging profiles of third generation photocaged azanucleosides **84** and **93**

It was observed that the 5'-caged analogue **84** had a notably slower deprotection rate compared to the 3'-caged analogue **93**. Illumination of **84** resulted in 38 and 18% of the caged analogue remaining after 30 and 60 min, respectively. Uncaging of analogue **93** resulted in 8% and 2% of the intact analogue after 30 and 60 min of irradiation. The conversion to dAC was quantitative.

The reason for the slower deprotection rate of 5'-caged dAC is still elusive. The ability to fit a biphasic exponential curve (as opposed to a monophasic exponential curve) suggested that there is an additional mechanism going on during the course of the uncaging. However, there were no other peaks observed in HPLC apart from the uncaged dAC and the phototag, indicating that there are no other side visible reactions taking place.

Figure 2.12 shows an additional observation that confirms the unusual behaviour of 5'-caged analogue. When a 1:1 mixture of 3'- and 5'-DEACMOC-caged dAC was irradiated at $\lambda = 365$ nm, the corresponding HPLC traces showed that after 30 mins, the 3'-caged dAC peak decreased more significantly compared to the 5'-caged dAC peak. This suggests that 3'-caged dAC uncaged more rapidly than the 5'-caged dAC. These findings confirm that an additional mechanism occurs in 5'-caged analogue that slows down its deprotection rate.

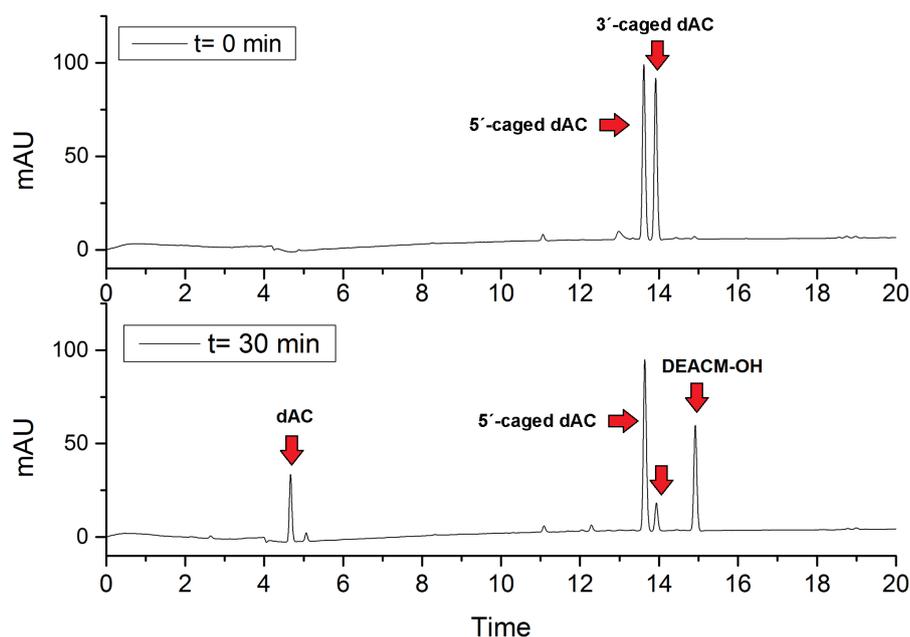
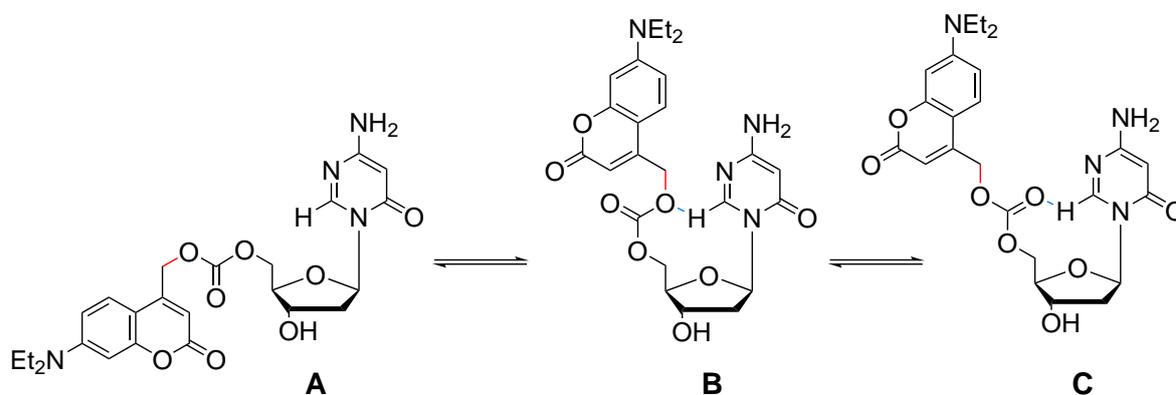


Fig. 2.12 HPLC analysis of photodeprotection of 1:1 5'- and 5''caged dAC mixture

A possible explanation for the slower uncaging rate of 5'-DEACMOC-dAC is proposed in Figure 2.20. In this, possible conformational isomers of 5'-DEACMOC-dAC are shown. The bond between C5 and O5 can freely rotate and as a result, higher energy isomers of the 5'-caged-dAC could form (B and C). Within these higher energy isomers, the phototag and the triazine ring come in close contact and stabilising interactions (such as hydrogen bonding, represented by the blue bond) could form between them. This stabilisation may interfere with the cleavage of the C–O bond (shown in red) and thus reduce photouncaging rate.



Scheme 2.20 Proposed stabilising interactions within 5'-DEACMOC-dAC 84

2.4.5 Stability studies

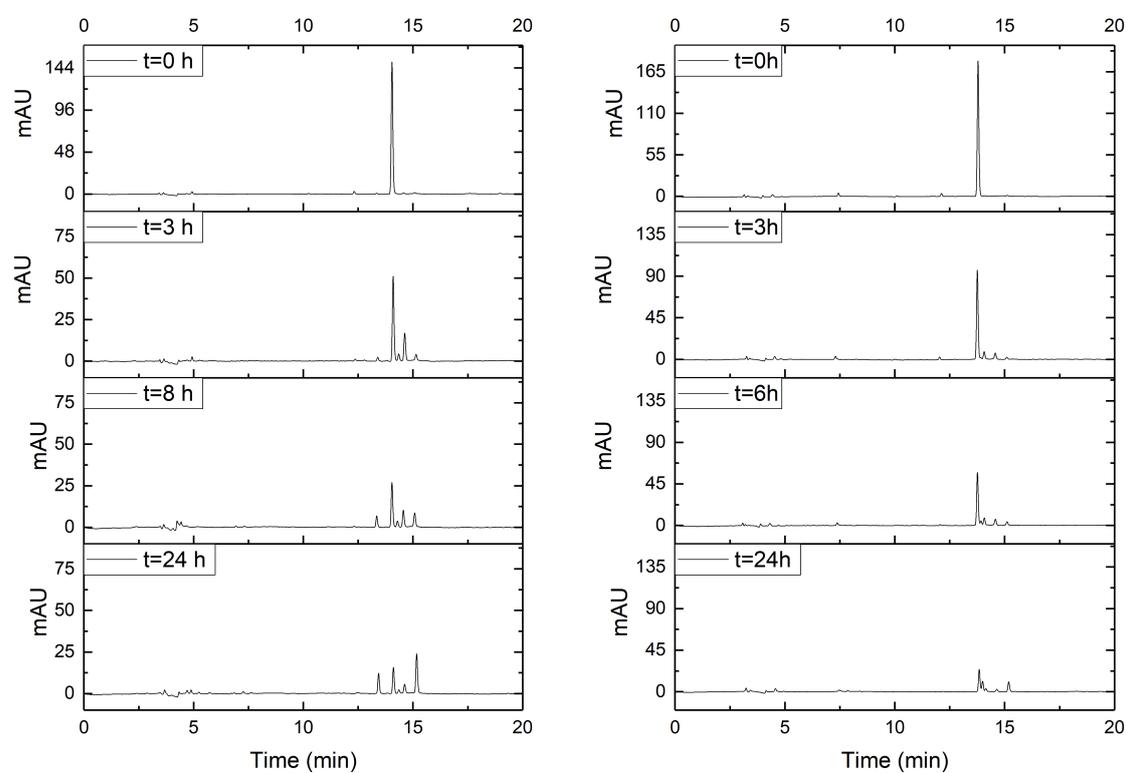
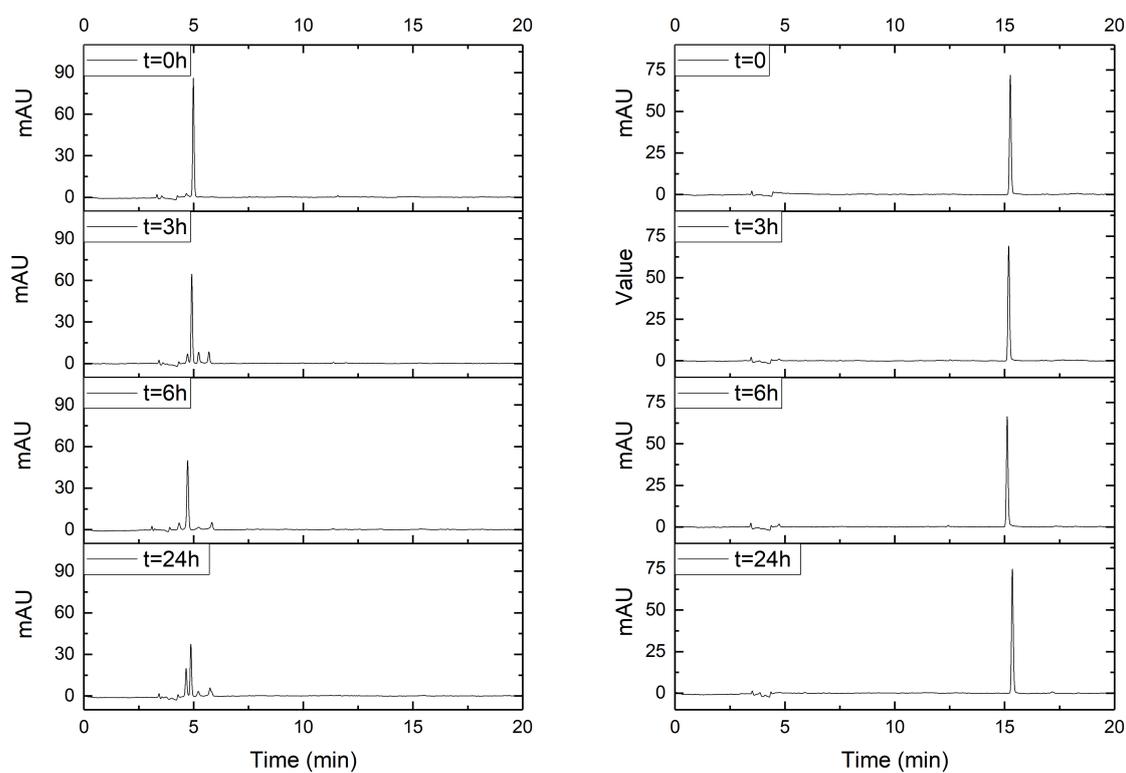
The motivation for synthesising the 3'- and 5'-caged dAC was to attain a higher hydrolytic stability compared to the NH₂ modified dAC. The compounds were analysed to test whether their stabilities were superior. The caged analogues were incubated in water at 37 °C and their stabilities were monitored *via* HPLC over 24 hours (Figure 2.13). Stability studies for dAC and DEACM-OH were also conducted.

3'-DEACMOC-dAC **93** had significantly improved stability compared to the N-DEACMOC-dAC **78** (Figure 2.13A). After 3 hours of incubation, 50% of the initial 3'-caged dAC peak was still intact as opposed to the completely degraded N-caged dAC **78** under the same conditions (Figure 2.6). In addition, no dAC was released as suggested by the lack of degradation peaks at $R_t = 5$ mins, indicating that the dark stability was markedly improved. However, the triazine ring of the 3'-caged dAC analogue likely underwent ring opening, as implied by the appearance of smaller degradation peaks between $R_t = 12$ and 15 mins.

The stability profile of 5'-DEACMOC-dAC **84** also suggested improved stability as no dAC was being released (Figure 2.13B). No alternative degradation peaks were observed, however the main caged-dAC peak significantly reduced in intensity after 24 hours of incubation.

Figure 2.13C showed that dAC **6** also underwent moderate degradation. The initial dAC peak at $R_t = 5$ mins reduced in intensity and gave rise to neighbouring degradation peaks, which is due to the triazine ring opening.¹⁵⁵ This is potentially the same process that leads to the moderate degradations of **84** and **93**. By contrast, the stability of DEACM-OH was unaffected after 24 hours of incubation (Figure 2.13D).

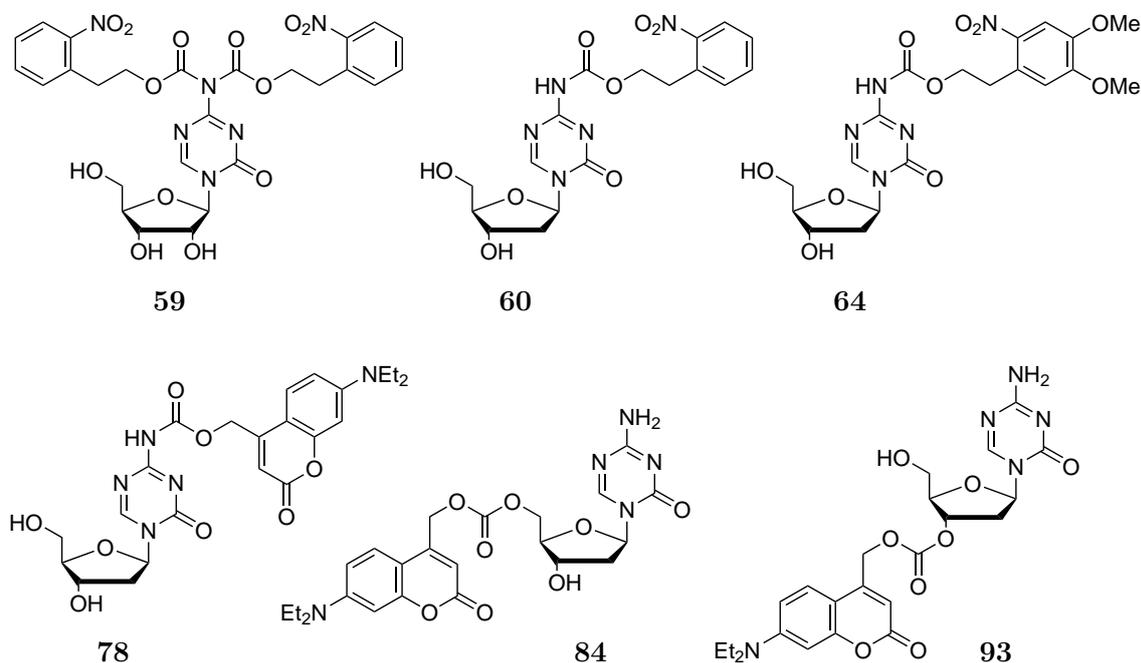
In summary, the HPLC stability studies indicated that the 3'- and 5'-caged analogues **93** and **84** had improved stabilities compared to N-caged dAC **78**. No release of dAC was observed during the incubation of these analogues at 37 °C in the dark, suggesting that the carbonate link between dAC and the phototag remained intact. However, the photocaged analogues did undergo moderate degradation during these incubations that could be attributed to the triazine ring-opening that is also observed in the parent dAC compound.

(A) 3'-DEACMOC-dAC, **93**(B) 5'-DEACMOC-dAC, **84**(C) dAC, **6**(D) DEACM-OH, **76**Fig. 2.13 Stability profiles of dAC, DEACM and analogues **93** and **84** at 37°C

2.5 Summary

In this chapter, the design and synthesis of six novel photocaged azanucleoside analogues have been presented and discussed (Figure 2.21). In general, dAC analogues had better stability compared to the AC analogue **59**. Furthermore, the coumarin-caged dACs showed superior photochemical characteristics compared to the NPE-caged analogues. It was also established that the positioning of the phototag had a strong effect on compound stability and photouncaging kinetics:

- Placing the phototag on the nucleobase (**78**) was not suitable due to the hydrolytic instability of the carbamate, likely caused by the triazine ring.
- Conjugation to the 5'-OH group (**84**) seemed to significantly improve dark stability of the caged compound, but the kinetic uncaging rate reduced possibly due to the stabilising interactions between the phototag and nucleobase.
- Photocaging on the 3'-OH gave **93** with fast uncaging kinetics and improved stability. In summary, this caging position was found to be the most suitable.



Scheme 2.21 Structures of novel photocaged azanucleoside analogues

For subsequent biological studies, we selected compound **93**, given its equivalent stability to dAC and fast deprotection rate under moderate, cell-compatible illumination.

Chapter 3

Bioactivity of Photocaged Azanucleosides

3.1 Introduction

Abnormal methylation such as gene promoter hypermethylation is a commonly observed molecular lesion in cancer cells (Figure 3.1A). DNMT enzymes are responsible for maintenance of methylation patterns in dividing cells. In order to lower methylation levels, DNMTs can be blocked with azanucleosides to achieve global demethylation (Figure 3.1B).^{111,141} The aim of the project was to achieve light-targeted demethylation using photocaged azanucleosides (Figure 3.1C). The objective was to test whether placing a phototag on an azanucleoside would block its bioactivity and result in unchanged methylation levels. It was also investigated whether light illumination can lead to the removal of the phototag and the release of the active azanucleoside, resulting in lowered methylation levels at targeted locations.

As described in the previous chapter, photocaged azanucleoside analogues have been successfully synthesised and characterised. In this chapter, the biological evaluation of the most promising caged analogue - 3'-DEACMOC-caged dAC **93** - is discussed in more detail. In the cell biological experiments, the effect of 3'-caged dAC (3'-CD) on DNA methylation levels, DNMT1 levels and cell viability were assessed. The human osteosarcoma (Saos-2) cell line was chosen as model system due to its high baseline methylation levels. First, currently used DNA Methylation assays in literature will be reviewed.

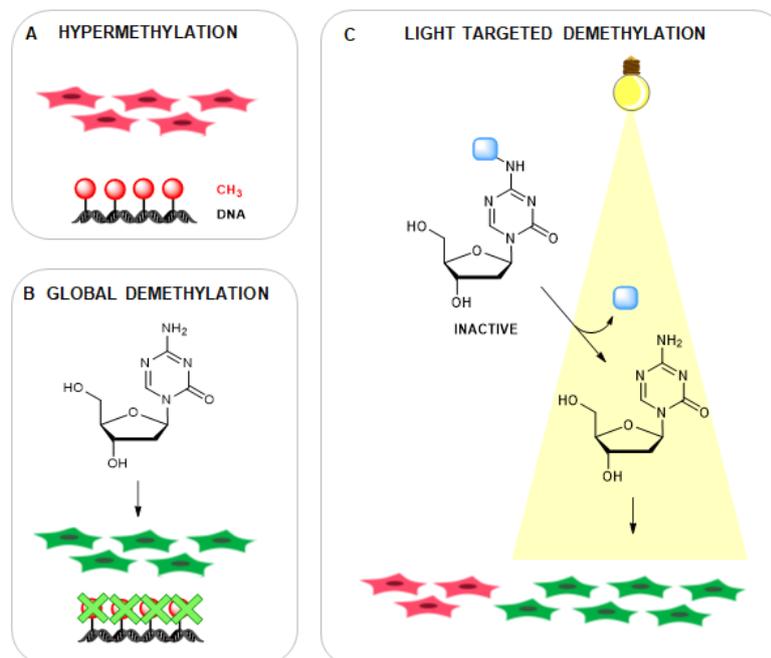


Fig. 3.1 Light tunable bioactivity of photocaged azanucleosides

3.2 DNA methylation assays in literature

3.2.1 Importance of DNA methylation analysis techniques

Natural changes in methylation levels can occur throughout cellular differentiation and organismal development.²²⁶ By contrast, abnormal methylation changes are highly associated with tumorigenesis and a number of other diseases. The analysis and profiling of DNA methylation are thus vital to our understanding of the role of epigenetic changes on cell cycle regulation and disease progression.^{227,228} DNA methylation profiling techniques have been extensively reviewed in literature.²²⁹ They can be classified into two main categories: locus-specific and global methylation analysis approaches.

3.2.2 Locus-specific DNA methylation analysis techniques

Locus-specific approaches are used to detect single methylation changes in the genome. Historically, these were one of the earliest methods used and typically involved the use of restriction endonucleases coupled to read-out with gel electrophoresis and hybridization

on southern blot.²³⁰ They were prone to false-positive results caused by incomplete digestion.

A further development of locus-specific approaches are the genome-wide methylation analysis techniques, where multiple methylation sites can be interrogated in parallel. These techniques are less labour intensive and have recently been coupled to high-throughput platforms such as microarrays or next generation sequencing (NGS). A summary of locus-specific and genome-wide methylation analysis techniques is shown in Table 3.1.²³¹

Pretreatment	Analytical step			
	Locus-specific analysis	Gel-based analysis	Array-based analysis	NGS-based analysis
Enzyme digestion	HpaII-PCR	Southern blot	DMH	Methyl-Seq
		RLGS	MCAM	MCA-seq
		MS-AP-PCR	HELP	HELP-seq
			MethylScope CHARM MMASS	MSCC
Affinity enrichment	MeDIP-PCR		MeDIP	MeDip-seq
			mDIP	MIRA-seq
			mCIP	
			MIRA	
Sodium bisulphite	MethyLight EpiTYPER Pyrosequencing	Sanger BS	BiMP	RBBS
		MSP	GoldenGate	BS-Seq
		MS-SNuPe	Infinium	BSPP
		COBRA		WGSBS

Table 3.1 Summary of locus-specific and genome-wide methylation analysis techniques²³¹

In general, the DNA sample requires pretreatment in order to differentiate methylated cytosines from their unmethylated counterparts. The pretreatments include enzyme digestion, affinity enrichment or sodium bisulfite treatment. The pretreated DNA can then be 'read' using a range of analytical techniques such as gel-, array- and next generation sequencing (NGS)-based analyses. The latter technique is more flexible and powerful as it allows for allele-specific DNA methylation analysis and can cover more of the genome with less input DNA. The disadvantage of these technologies is the high cost and thus the limited number of samples that can be analysed. In addition, these technologies require specialised equipment.

3.2.3 Global methylation analysis techniques

Global methylation analysis is the other category of DNA methylation profiling techniques. Using these techniques, the overall methylation levels of cells can be measured without focusing on the individual methylation sites. Measurements from global DNA methylation profiling can be used for classifying and characterising cancers and other diseases.²³² In addition, global methylation profiling can aid the identification of biomarkers and can be useful for drug screening.²²⁹

HPLC

For many years, the method of choice for quantitative determination of the global DNA methylation levels has been high performance liquid chromatography (HPLC).²³³ Similar to the locus-specific and genome-wide techniques, the DNA is subjected to pretreatment whereby the total genomic DNA is hydrolysed to deoxyribonucleotides using a combination of nuclease enzymes (deoxyribonuclease I and nuclease P1) followed by further treatment with alkaline phosphatase to produce deoxyribonucleosides.²³⁴ The products are then separated by standard reverse-phase HPLC and quantified *via* monitoring the UV absorbance at $\lambda = 254$ nm. The limitation of this technique is that it requires relatively large amounts of DNA (3-10 μ g).

HPCE

High performance capillary electrophoresis (HPCE) is a more simple, rapid and cost-effective alternative of HPLC.²³⁵ In this technique, the digested DNA sample mixture is separated in an open-tube capillary system upon application of electric current.²³⁶ Generally, the sensitivity of detection is high, but not enough for routine processing of tissue specimens.

LC-MS/MS

In order to improve the detection sensitivity of HPLC techniques, new methods have been developed by combining HPLC separation with mass spectrometry analysis. Liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) requires smaller quantities of hydrolysed DNA (50-100 ng). In mammalian DNA 2-5% of cytosines are methylated. LC-MS/MS is able to detect ranges between 0.05 and 10%

and it can also confidently detect 0.25% differences between samples. This method has many advantages, however it also requires specialist equipment that is less abundant than HPLC.

ELISA-based assays

Global methylation can also be measured by immunofluorescence using antibodies. Antibody-based assays are more affordable alternatives to the previously mentioned techniques. These methods rely on anti-5mC antibodies to selectively detect methylated cytosines by quantitative fluorescence in an ELISA-like reaction. The final readout in fluorescence or colorimetric change is proportional to the %5mC present in the sample DNA. The advantages of this technique is that they are quick to perform (several commercial kits are available) and they serve well to detect large changes in global methylation. The limitation is that these are only suitable for rough estimation of DNA methylation and are also prone to high variability.

LINE-1 Pyrosequencing

Another global methylation technique involves the determination of the methylation states of certain repetitive regions such as *Arthrobacter luteus* (ALU) or long interspersed nuclear elements 1 (LINE-1).²²⁹ Repetitive element methylation is directly proportional to global methylation content. The genomic DNA is subjected to bisulphite treatment followed by PCR amplification of ALU or LINE-1 consensus sequences. The methylation of the amplified fragments is then quantified by pyrosequencing. This method is suitable for medium high throughput analysis of samples.

LUMA Assays

The most recently developed technique is the luminometric methylation assay (LUMA), first reported by Karimi *et al.* in 2006.²³⁷ This technique enables researchers to study global methylation by using methylation-sensitive restriction enzymes followed by pyrosequencing. The specificity of the method is high and the variability is low. These are essential for the precise measurement of changes in global methylation. The advantage is that only small amounts of DNA are needed (250-500 ng), however high quality DNA is essential to ensure that complete enzymatic digestion occurs.

3.3 Selecting a suitable DNA methylation assay

The method of choice should deliver an unbiased answer to the biological question being asked.²³⁸ Some of the key considerations are listed below:

- Amount and quality of DNA samples
- Sensitivity and specificity requirements of the study
- Robustness and simplicity of the method
- Availability of bioinformatics software (for selected techniques)
- Availability of specialized equipment and reagents
- Cost

In this project, DNA methylation assays will be used to determine the effect of dAC analogues on baseline methylation levels. dAC is a global demethylator. By inhibiting DNMT enzymes, it can indirectly remove methylation marks across the genome in a non-specific manner. Methylation changes induced by dAC are generally quantified by global DNA methylation profiling techniques such as HPCE,^{111,113} LINE-1¹⁶⁶ or LUMA.¹⁴⁶ Locus-specific and genome-wide analysis of methylation changes are also used to assess gene-specific responses to azanucleoside treatments.²³⁹

The question of this thesis is whether photocaged versions of dAC can achieve light-dependent demethylation in cancer cells. Despite the detailed information that could be extracted from genome-wide sequencing studies, our primary goal was to measure the global methylation changes in response to photocaged dAC treatment to assess whether the phototag can successfully block the biological effects of dAC.

The optimal DNA methylation assay for our study would allow the simultaneous and high-throughput analysis of multiple samples (>100) to enable the testing of different treatment conditions (concentrations, light, in triplicate). A sensitive yet simple method was required due to the limited availability of specialized equipment. Consequently, two methylation assays (ELISA-based assays and LC-MS/MS methods) were trialled to assess their suitability for our study purposes.

3.3.1 Trial: ELISA assays

ELISA-based assays were initially chosen due to their simplicity, high speed and low cost. Prior to conducting trial runs, the shortcomings and limitations of ELISA-based assays for measuring DNA methylation changes were carefully considered.

As mentioned before, ELISA-based assays are suitable for a rough estimation of DNA methylation, however its ability to be able to detect small changes in methylation levels has been questioned by others.²³⁸ It has been suggested that ELISA can only detect relatively large changes in methylation levels (e.g. 1.5-2x fold) due to the high level of inter and intra-assay variabilities.²⁴⁰ These variabilities can arise due to a number of factors such as human technical errors (e.g. pipetting inaccuracies, splashing of reagents between wells, variability due to freeze-thaw cycles), machine errors (e.g. uncalibrated pipettes, plate readers and softwares etc.) and contaminations of plates, samples and reagents. For a reliable reading, it is important that the % coefficient of variability is less than 15% for inter-assay (between independent assays) and less than 10% for intra-assay (variability of sample readings within the same assay). The other limiting factor of ELISA-based assays is their high signal-to-noise ratio. The average methylation level of human DNA ranges between 2-6%.³² This measurement represents the % of 5-methylcytosine in the DNA present compared to total methylated and unmethylated cytosines. Therefore, it is crucial for the ELISA-based assay to be able to detect slight changes within this range in order to be suitable for our applications.

In 2012, Kremer and colleagues described a novel methDNA-ELISA system that was able to accurately quantify global DNA methylation levels in the regions between 1-10% making it highly suitable for the typical ranges from 2-6% in mammalian genomes.²⁴¹ In another work by So and colleagues, the changes in methylation levels of BEAS-2B (human lung) cells were successfully shown in response to 24 hour azacytidine treatment using Epigentek's 5mC ELISA-based assay.²⁴⁰ Given the high frequency of publications that mentioned the use of Epigentek's ELISA assay (258 as of Aug 2017), the MethylFlash Methylated DNA 5mC Colorimetric Quantification Kit (Epigentek) was chosen as our initial tool for trial methylation measurements.

The assay comprised a 48-well plate allowing the measurement of 12 control and 36 test samples. The 12 control samples included 2 negative (unmethylated DNA) and 10 positive controls (DNA with increasing amounts of 5mC) in order to establish a calibration curve and validate the assay. The details of the protocol are discussed in the Experimental Section. In addition to the control samples, 36 test genomic

DNA samples were included. These were extracted from Saos-2 cells that were treated with increasing concentrations of dAC (0-365 μM) in duplicate. Figure 3.2 shows the measured average absorption values at $\lambda = 450 \text{ nm}$ (OD450 nm) that are proportional to the amount of 5mC detected in the DNA samples.

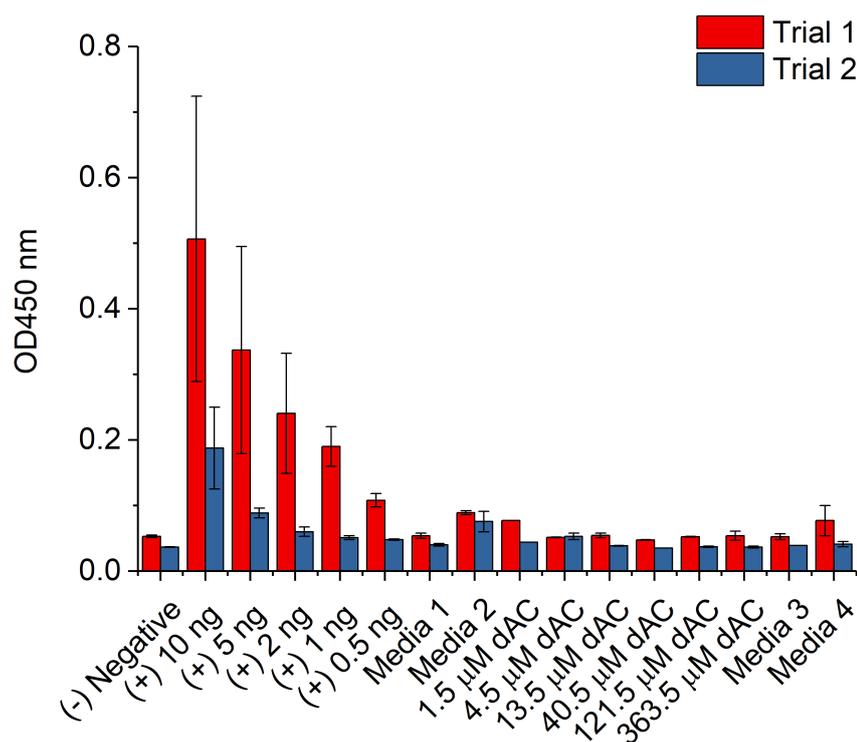


Fig. 3.2 Results of ELISA-based assay trials for assessing methylation changes in SaOS-2 cells. The graph depicts optical density measurements at $\lambda = 450 \text{ nm}$ from two independently conducted assays, Trial 1 (red) and Trial 2 (in blue). The optical density measurements are proportional to the amount of 5 mC detected in the DNA samples. For each trial, 6 control and 10 test samples were tested in duplicates. The control samples included negative and positive samples containing 0.5, 1, 2, 5, 10 ng methylated DNA (included in the test kit). The test samples were extracts of SaOS-2 cells that were previously treated with increasing concentrations of dAC (0-365 μM) for 24 hours. The results show that ELISA-based assays would not be suitable for future methylation measurements due to lack of sensitivity and high level of variability in the results obtained.

The first limitation of the assay was the measurement of samples in duplicates instead of triplicates. Ideally, samples should have been done in triplicates, but for the interest of screening more dAC treatment concentrations, we decided to lower the power of the samples for an initial screen, after which a more thorough screen could have been

conducted with triplicate samples and narrower treatment concentrations. However, the initial results themselves seemed problematic.

The results highlighted two problems with this technique. Firstly, the measured absorption values for positive control samples were not reproducible between different trial runs or even between replicates of the same run (represented by the large error bars). Secondly, the measured values for the genomic DNA samples were extremely low, equivalent to the negative control. This suggests that either the concentrations of DNA samples were not sufficient or the DNA samples had very low methylation levels. The insufficient DNA concentration was likely an explanation in this case. However, the reproducibility issue remained a more significant concern. Various changes were made in the protocol to improve the reproducibility such as increasing DNA binding time, development time or keeping the reagents at room temperature. Unfortunately, none of these attempts were successful. Due to the unsuitability of this approach to our studies, with high variability of the results both within an assay and between assays, the ELISA-based assays were not pursued further.

3.3.2 Trial: LC-MS/MS analysis

LC-MS/MS was used as an alternative method to determine methylation levels. LC-MS/MS analysis has reportedly high sensitivity and accuracy rates in determining % 5mC values in genomic DNA samples.²⁴²

For the LC-MS/MS equipment and technical expertise, a collaboration with Professor Shankar Balasubramanian's group from Cambridge (UK) was set up. The LC-MS/MS analyses of genomic DNA samples were performed by Dr Shiqing Mao (Department of Chemistry, University of Cambridge).

Briefly, Saos-2 cells were treated with dAC every 24 h for 4 days at three different concentrations (0.5, 1 and 1.5 μM). Each treatment was carried out in triplicate. 24 h after the last treatment, cells were collected and the genomic DNA was extracted and submitted to Cambridge for LC-MS/MS analysis.

The preliminary results indicated that the measurements were highly reproducible. Figure 3.3 shows the measured % 5mC values in response to different concentrations of dAC treatments. The untreated cells had on average 2.4% 5mC. Upon treatment with 0.5 μM dAC, the methylation dropped to 1.3% (53% of baseline methylation) and further decreased to 0.74% and 0.75% in response to 1 and 1.5 μM dAC respectively.

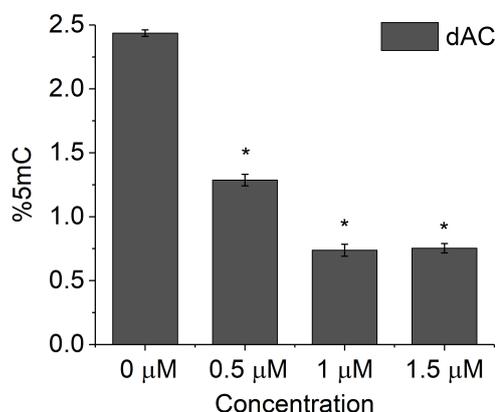


Fig. 3.3 Results of LC-MS/MS methylation assay trial (n=3). The graph depicts the measured % 5-mC content of DNA samples extracted from SaOS-2 cells that were treated with 4 different concentrations of dAC: 0, 0.5, 1 and 1.5 μM . The cells were treated with dAC every 24 h for 4 days and each treatment was carried out in triplicates. The results represent the mean \pm standard error of three independent samples per treatment. Asterisks above SD bars show the statistical significance of the difference with respect to control treatment (0 μM) after 4 days. * $p < 0.05$ (Student t-test). The corresponding p-values for the 0.5, 1 and 1.5 μM dAC treatments were 0.0017, 0.0004 and 0.0007 respectively.

The demethylating effect of dAC reached a plateau at 1 μM , potentially due to the passive demethylating nature of dAC that is limited by the cell replication cycle.

Given the promising and reproducible results from the LC-MS/MS analysis, we selected this approach as our main analytical technique to determine DNA methylation changes.

3.4 DNA methylation analysis *via* LC-MS/MS

3.4.1 Optimisation of treatment conditions

Treatment concentration, duration and frequency

For the following treatments with 3'-caged dAC (3'-CD), we aimed to establish an optimised treatment schedule that would minimise toxicity to the cells and maximise the amount of DNA that could be extracted. In the trial experiments, it was observed that daily treatment of the cells with the highest concentration of dAC (1.5 μM) had a significant effect on cell growth. Hence, the treatment duration was reduced to 24 h as it is known that dAC can induce effective demethylation after 24 h.²⁴³ In addition, lower treatment concentrations (0.1 and 0.3 μM) were also included.

Cell line screening

Our initial cell experiments were carried out on Saos-2 cells. Although these cells showed concentration-dependent decrease in methylation for the 4-day treatment schedule (Figure 3.3), when switching to the 24-h treatment schedule, the concentration-dependency was no longer visible.

In order to determine if this is a cell-line dependent phenomenon, additional cell lines were screened in order to evaluate their sensitivity to dAC. For the cell-line screening, human bladder (T24), human lung carcinoma (A549) and human embryonic kidney (HEK293T) cells were treated with increasing concentrations of dAC for 24 h and the genomic DNA was extracted and submitted for LC-MS/MS analysis in Cambridge.

The results show that among the cell lines, Saos-2 cells had the highest baseline methylation (3%), whereas the remaining cell lines had on average 2.5% 5mC (Figure 3.4). As discussed before, Saos-2 cells did not show a concentration dependent response to dAC treatment. At the lowest dAC concentration (0.1 μM), the methylation levels dropped to 2.2%, which remain largely unchanged for higher concentration treatments. This pattern was also the case for A549 cells. In contrast, T24 and HEK293T cells showed promising concentration dependent trends, and hence were chosen as additional control cell lines for further cell biological studies.

Note about the number of samples: One limitation of this assay is that the samples were not run in triplicates. Given the number of cell lines and treatment concentrations

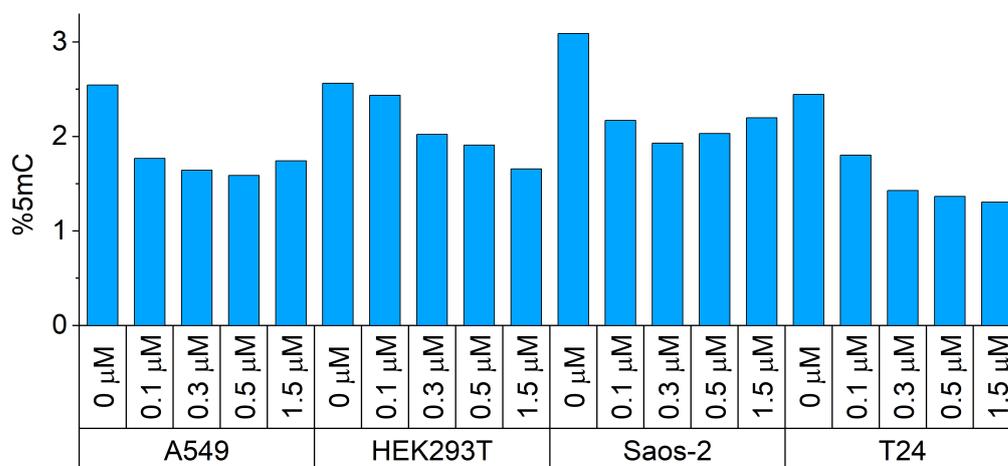


Fig. 3.4 **Cell line screening in response to dAC treatment (n=1)**. LC-MS/MS analysis of methylation changes in four different cell lines were conducted to measure their sensitivity to dAC treatment. A549, HEK293T, SaOS-2 and T24 cells were treated with 5 different concentrations of dAC (0, 0.1, 0.3, 0.5 and 1.5 μM) for 24 h. Among the cell lines tested, HEK293T and T24 cells showed the strongest concentration-dependent methylation changes in response to dAC treatments.

that were considered for this initial screen (20 conditions), the samples were run in singlicate to enable faster turnaround of results. Each time a treatment was conducted, the samples were sent to a different laboratory (in Cambridge), where the main limiting factor was the availability of the LC-MS/MS machine. Since the treatments of the chosen cell lines were repeated several times (and in triplicates) in subsequent experiments, it felt was justified to conduct the initial cell screening in singlicate.

3.4.2 Measurement of DNA methylation changes

After demonstrating that LC-MS/MS can accurately determine methylation levels of our DNA samples, Saos-2, T24 and HEK293T cells were treated with 3'-CD **93** for 24 h at different concentrations (0.1-1.5 μM). After treatment, the cells were either illuminated for 1 h at $\lambda = 365$ nm or were kept in the dark. dAC treatments were also included as controls. 24 h after treatment, the cells were collected and the isolated genomic DNA was submitted for LC-MS/MS analysis (carried out by Dr Shiqing Mao). All cells were treated with the 3'-CD **93**, but in the case of T24 cell lines, a treatment set with 5'-caged dAC (5'-CD) **84** was also included.

DNA methylation changes in Saos-2 Cells

Figure 3.5 presents the DNA methylation change (expressed as % of baseline methylation) in Saos-2 cells in response to 3'-CD and dAC treatments. Six different concentrations were administered, including the media only treatment at 0 μM . Three main observations were drawn from this dataset.

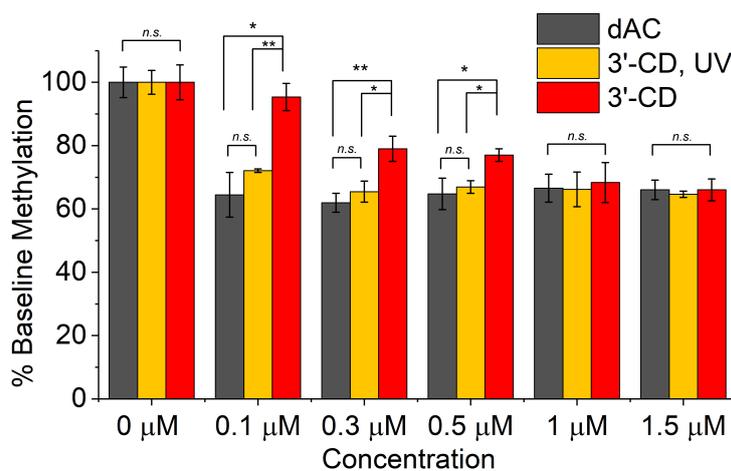


Fig. 3.5 LC-MS methylation analysis of SaOS-2 cells (n=3). In order to determine the light-dependent effect of 3'-CD on methylation levels, SaOS-2 cells were treated with 6 different concentrations of dAC and 3'-CD (0, 0.1, 0.3, 0.5, 1 and 1.5 μM) with and without UV irradiation. The light-dependent demethylating effect of 3'-CD was most significant at 0.1 μM . At 0.1 μM , 3'-CD maintained 95% \pm 4.3 % of the baseline methylation, whereas the same treatment followed by UV treatment resulted in a 23% drop in methylation ($p=0.0065$) to 72 \pm 0.5%, due to the uncaging and release of active dAC. This treatment achieved comparable methylation levels to 0.1 μM control dAC treatment (64 \pm 7%, $p=0.09$). The light dependent effect was also significant at 0.3 μM and 0.5 μM ($p=0.010$ and $p=0.0167$ respectively), however little differentiation was observed at 1 and 1.5 μM treatments ($p=0.524$ and $p=0.132$). This trend is due to the concentration dependent demethylating effect of 3'-CD in the absence of UV light. Treatments were carried out in triplicates. Key for symbols of statistical significance: ns: $p > 0.05$, *: $p < 0.05$, **: $p < 0.01$. Statistical significance was determined using a two-tailed Student's t-test and standard deviations. For all analyses, $p < 0.05$ was considered significant

Firstly, the Saos-2 cell line is highly sensitive to dAC as mentioned before. The lowest concentration treatment (0.1 μM) induced a 35% drop in the baseline methylation levels. This effect remained constant with increasing dAC concentrations.

Secondly, the UV illumination resulted in the successful activation and release of dAC. This is confirmed by the equivalent methylation levels of dAC treated and 3'-CD treated and illuminated (CD UV) cells.

The final point is that the non-illuminated, 3'-CD treated cells also resulted in decreased methylation levels especially at higher treatment concentrations. Methylation dropped to 78%, 77%, 68% and 66% in response to 0.3, 0.5, 1 and 1.5 μM 3'-CD respectively. However at 0.1 μM 3'-CD, the methylation levels of the cells remained unchanged.

These results suggest that the light-dependent demethylating effect of 3'-CD was most notable at lower concentration treatments (0.1 μM). The phototag sufficiently masked the bioactivity of dAC, resulting in an unchanged methylation level. Upon irradiation, the active dAC was successfully released and achieved equivalent methylation drop compared to the non-caged dAC control treatment. Unfortunately at increasing treatment concentrations, significantly increasing activity was also observed even in the absence of light.

DNA methylation changes in T24 Cells

Figure 3.6A shows the DNA methylation changes of T24 cells in response to 3'-CD **93** and dAC. The number of treatment concentrations were reduced to four (0, 0.1, 0.5 and 1.5 μM). In addition, T24 cells were also treated with 5'-CD **84** in a separate set of experiments to investigate whether different analogues of photocaged dAC exert different biological effects (Figure 3.6B).

T24 cells exhibited a concentration-dependent response to dAC treatment. Similarly to Saos-2 cells, 0.1 μM dAC treatment resulted in a 35% drop in baseline methylation. Increasing the concentration to 0.5 μM led to a further drop of 17%, which remained unchanged at 1.5 μM . For the non-illuminated 3'-CD treatments, a concentration-dependent reduction in methylation levels was also detected, with the strongest effect observed at 1.5 μM (34% drop). As 3'-CD is expected to stay inactive in the dark and maintain methylation levels, this pattern indicates unexpected and unwanted activity. This trend of 'dark activity' was also observed in Saos-2 cells.

Upon light exposure, the baseline methylation dropped to 81%, 57% and 51% in response to increasing 3'-CD concentrations. This confirmed that light-dependent release of dAC was successful. However, it was noted that light-activation was less efficient at lower treatment concentrations. 0.1 μM of the illuminated 3'-CD lowered

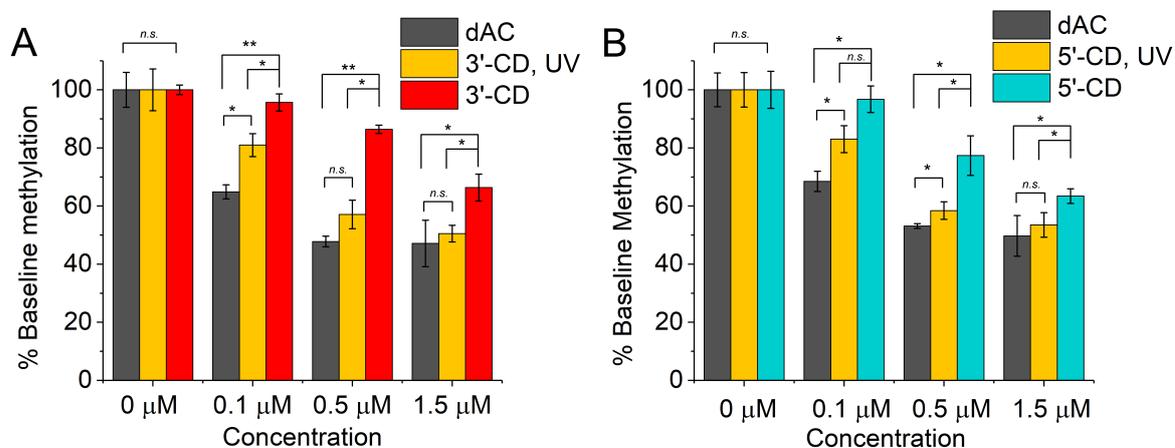


Fig. 3.6 LC-MS methylation analysis of T24 cells in response to (A) 3'-CD treatment and (B) 5'-CD treatment (n=3). In two separate experiments, T24 cells were treated with dAC and either 3'-CD or 5'-CD with or without UV illumination (for 1 hour) at 4 different concentrations (0, 0.1, 0.5 and 1.5 μM). (A) **3'-CD treatment set.** The light-dependent demethylating effect of 3'-CD has been demonstrated across all concentrations with varying extent. In response to 0.1 μM 3'-CD in the dark, methylation was maintained at 95% ± 0.1% of the baseline. The same treatment followed by irradiation resulted in a significant drop to 81% ± 0.1% (p=0.0304). In comparison, the control 0.1 μM dAC achieved a methylation drop to 65% of the baseline. This suggested that the light activation of 0.1 μM 3'-CD may not have released equivalent amounts of uncaged dAC compared to control dAC. On the other hand, it was shown that at 0.5 μM and 1.5 μM, the light-activated 3'-CD achieved significant drop in methylation levels (p=0.0017 and p=0.0219 respectively) and they were also comparable to the dAC controls of the same concentration (p=0.091 and 0.6766). (B) **5'-CD treatment set.** Addition of 0.1 μM 5'-CD did not result in statistically significant changes in methylation compared to baseline (96.7% ± 4.5%), and although irradiation after the same treatment did lower methylation to 83% ± 4.6%, this was not deemed statistically significant (p=0.0994). At 0.5 μM and 1.5 μM, light activation resulted in significant reduction of methylation. At 0.5 μM, methylation dropped from 77.4 ± 6.8 % to 58.4 ± 3 % (p=0.0418) and at 1.5 μM methylation reduced from 63.4 ± 3.5% to 53.5 ± 5.2% (p=0.0245). Methylation levels equivalent to control dAC were only achieved at 1.5 μM treatment (p=0.121 suggesting that the two values are not significantly different). In conclusion, the light activating effects of 5'-CD were moderate considering the weak light activating effects at lower concentration. Treatments were carried out in triplicates. Key for symbols of statistical significance: ns: p > 0.05, *: p < 0.05, **: p < 0.01. Statistical significance was determined using a two-tailed Student's t-test and standard deviations. For all analyses, p < 0.05 was considered significant

methylation levels to 81%, as opposed to the 65% achieved by 0.1 μM dAC. For the 5'-CD treated cells, the pattern was almost identical to its 3'-caged counterpart (Figure 3.6B).

DNA methylation changes in HEK293T Cells

Figure 3.7 shows the LC-MS methylation analysis results for HEK293T cells. In general, the methylation levels of HEK293T cells were less sensitive to treatments with dAC and 3'-CD. At 0.1 μM , HEK293T cells showed almost no response, with nearly unchanged DNA methylation levels. At 0.5 and 1.5 μM treatment concentrations, dAC decreased methylation levels to 72% and 63% of baseline methylation, respectively.

By contrast, 3'-CD did not induce any response at 0.5 μM , and only a moderate decrease at 1.5 μM treatments (75% for irradiated 3'-CD and 85% for control 3'-CD treatments). The limited activity of irradiated 3'-CD treatments could be due to their reduced transport into HEK293T cells, interference with light-activation or other sources of compound inactivation within the cells.

The growth rate of HEK293T cells is significantly lower than that of Saos-2 and T24 cells. Their doubling time is 24 h, whereas for Saos-2 and T24 cells, the doubling times are 19 and 17 h respectively. Although the cells could only complete one replication cycle, considerable demethylation occurred at 0.5 and 1 μM dAC treatments. This could be due to the additional replication-independent effects of dAC, described by Ghoshal and colleagues.^{117,118}

3.4.3 Summary of LC-MS/MS methylation results

The LC-MS/MS analysis provided valuable insights into the effect of dAC and 3'-CD on the methylation levels of three different cell lines (Saos-2, T24 and HEK293T). In general, all cell lines responded to dAC treatments, though with varying sensitivities in the following order: Saos-2 > T24 > HEK293T.

The low sensitivity of HEK293T cells could be due to their lower growth rates, which limit the amount of dAC that could be incorporated into the replicating DNA. HEK293T was also the only non-cancerous cell line in our studies. The origin of this cell line (human embryonic kidney cells) may also be accountable for the sensitivity differences observed.

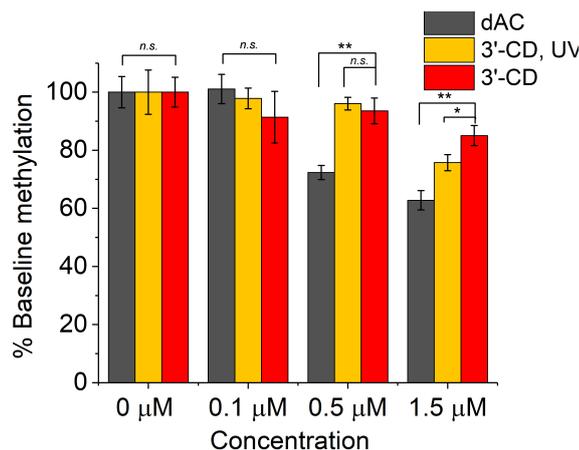


Fig. 3.7 LC-MS methylation analysis of HEK293T cells (n=3). In order to determine the light-dependent effect of 3'-CD on methylation levels on HEK293T cells, dAC and 3'-CD at 4 different concentrations (0, 0.1, 0.5 and 1.5 μM) were administered with or without UV light. In general, HEK293T cells were less sensitive to both dAC and 3'-CD. At 0.1 μM , none of the treatments resulted in statistically significant methylation changes compared to baseline. At 0.5 μM , dAC reduced baseline methylation to $72 \pm 3\%$, however 3'-CD treatment resulted in almost unchanged methylation levels with or without illumination. HEK293T cells responded to the 1.5 μM treatment series with the greatest extent. Control dAC achieved $62 \pm 4\%$ of baseline methylation, whereas 3'-CD treatment resulted in $85 \pm 4\%$ methylation which was further reduced to $75 \pm 3\%$ upon illumination, which was statistically significant ($p=0.0484$). In conclusion, HEK293T cells showed low sensitivity to dAC and caged-dAC treatments and in addition the extent of light-sensitivity was milder compared to other cell lines. Treatments were carried out in triplicates. Key for symbols of statistical significance: ns: $p > 0.05$, *: $p < 0.05$, **: $p < 0.01$.

Our results also show that with 3'-CD light-dependent demethylation was successfully achieved in Saos-2 and T24 cell lines, but only at low concentrations ($<0.5\mu\text{M}$). Remarkably, demethylation could be induced upon illumination to achieve a biological effect nearly equivalent to dAC's.

Unfortunately, demethylating activity in response to non-illuminated 3'-CD was also observed in all cell lines tested. The 'dark activities' were more significant at higher treatment concentrations ($>0.5\mu\text{M}$) and notably in Saos-2 cells. This effect could be due to chemical or enzymatic hydrolysis of the photocaged analogue (3'-CD) in the dark that would lead to the release of active dAC. Since Saos-2 cells are highly sensitive to the lowest treatment concentrations of dAC, the early onset of dark activity (starting at $0.3\mu\text{M}$) could be explained by this characteristic.

Another possible explanation for the unexpected activity is an alternative mechanism induced by the 3'-CD that does not proceed *via* the conventional replication-dependent mechanism of dAC. Ghoshal and colleagues reported an alternative pathway whereby dAC may interact with other enzymes such as δPKC in order to exert methylation changes *via* DNMT1 degradation.¹¹⁷ The exact details of this mechanism are still unknown. It was possible that 3'-CD could exert bioactivity through this pathway in the absence of light activation. Intrigued by their findings, the next aim was to investigate the DNMT1 levels in response to 3'-CD treatment.

3.5 Western blot analysis

The next objective was to determine whether the reduction in methylation was also detectable in terms of lower DNMT1 levels. DNMT1 enzymes are endogenously expressed in most cell lines and are responsible for maintaining methylation patterns. The changing levels of DNMT1 can be detected using Western blot analysis. Degradation of DNMT1 can be induced by azanucleosides and this has been studied in a number of cell lines including lymphoma, HeLa and HCT116 cells.^{114,117}

We were particularly interested in the sensitivity of DNMT1 levels to different concentrations dAC and 3'-CD. Saos-2 and T24 cells were treated under the same conditions as described in the methylation analysis studies. In addition to 3'-CD and dAC, 5'-CD was also included as treatments for comparison. Whole cell extracts from treated Saos-2 and T24 cells were subjected to Western blot analysis to assess DNMT1 levels.

3.5.1 Western blot analysis of Saos-2 cell extracts

Western blot analyses (Figure 3.8) show the expression levels of DNMT1 and control glyceraldehyde 3-phosphate dehydrogenase (GAPDH) enzymes of Saos-2 cells in response to dAC, 3'-CD and 5'-CD with and without illumination ('UV' or 'CONT').

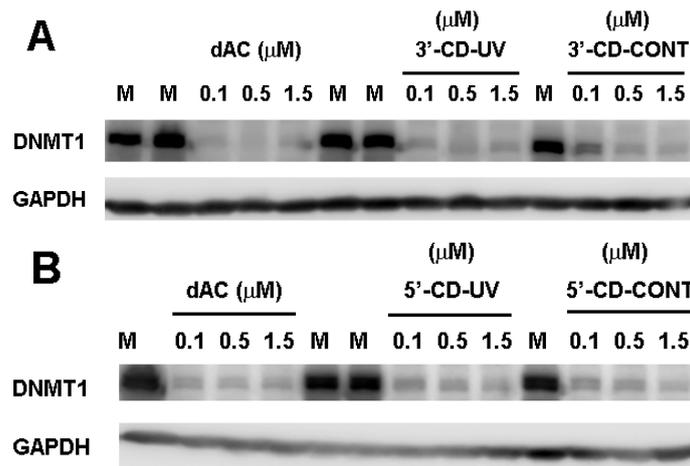


Fig. 3.8 Western blot analysis of SaOS-2 cell extracts. In order to determine the effect of 3'-CD and 5'-CD treatment on the cellular DNMT1 levels, western blot analysis was carried out on treated SaOS-2 extracts. SaOS-2 cells were treated with 4 concentrations (0, 0.1, 0.5, 1.5 μM) of dAC and 3'-CD (with and without UV, annotated as 3'-CD-CONT and 3'-CD-UV respectively) for 24 h, followed by extraction of cell lysate (see Experimental Section). Western blot analysis was performed with anti-DNMT1 and anti-GAPDH. Untreated (M) SaOS-2 cells have high level of baseline DNMT1 represented by the thick bands. No treatment followed by UV irradiation also did not have an effect on DNMT1 protein levels. On the other hand, dAC treatment even at the lowest concentration. The effect was similar looking at 3'-CD-UV and 5'-CD-UV treatments. On the other hand, the 3'-CD-CONT treatments maintained a significant level of DNMT1 at 0.1 μM , which can still be detected at 0.5 μM treatments, but not at 1.5 μM . This effect can also be observed for 5'-CD-CONT but to a lower extent.

Non-treated cell extracts (M) had high DNMT1 expression levels, represented by thick bands on both western blots, which were not affected by UV illumination. However, under all other treatment conditions the DNMT1 levels were significantly diminished in both treatment sets, but to a slightly smaller extent for 3'-CD control at 0.1 μM .

The faint band observed at 0.1 μM indicates that the phototag partially blocked the bioactivity of dAC, resulting in a limited degradation of DNMT1. This is in line with expectations. At higher concentration treatments, this effect was not observed.

3.5.2 Western blot analysis of T24 cell extracts

Western blots of T24 cell extracts show similar patterns to those of Saos-2 cells (Figure 3.9). Untreated T24 cells had high expression levels for DNMT1 that were not altered by UV illumination. dAC and 3'-CD and 5'-CD treatments led to complete removal of DNMT1 levels, with the exception of the lowest concentration of non-illuminated 0.1 μM 5'-CD (5'-CD-CONT). Similarly to Saos-2 cells, these observations are in line with expectations. In non-illuminated photocages dAC treatments (3'-CD-CONT and 5'-CD-CONT), considerable levels of DNMT1 were detected.

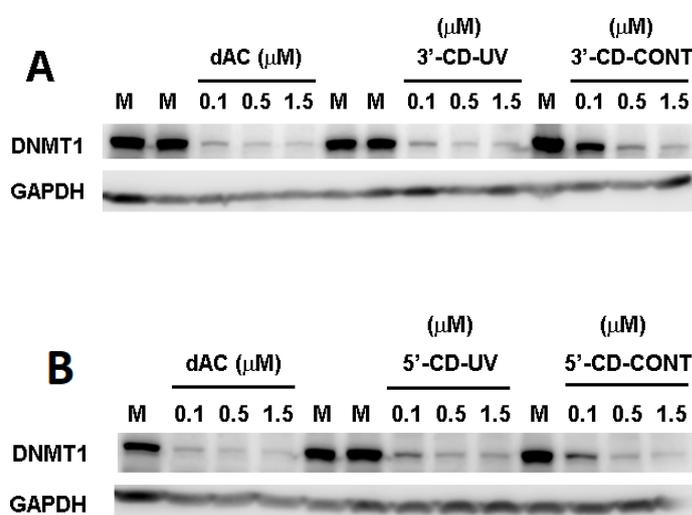


Fig. 3.9 Western blot analysis of T24 cell extracts. The effect of caged-dAC treatment on DNMT1 levels in T24 cells were also investigated via Western blot analyses. T24 cells were treated with 4 concentrations of dAC, 3'-CD and 5'-CD and were either kept in the dark ("CONT") or were irradiated for 1 h ("UV"), followed by 24 h incubation and extraction of the cell lysate (see Experimental Section). Similarly to the SaOS-2 cell extracts, untreated T24 cells had high level of DNMT1 indicated by the thick bands (M). dAC and irradiated 3'-CD, 5'-CD treatments had a significant effect on DNMT1 levels, resulting in the complete loss of DNMT1 band at all dAC treatment concentrations and the higher 3'-CD and 5'-CD treatment concentrations (0.5 and 1.5 μM). At the lowest irradiated treatment concentration, a weak band can be observed for both 3'-CD and 5'-CD, suggesting that a very small amount of DNMT1 is still present. DNMT1 levels were higher when cells were treated with either 3'-CD or 5'-CD and were not subjected to irradiation, especially at lower treatment concentrations (0.1 μM).

3.5.3 Summary of western blot results

In this section, the DNMT1 protein levels of Saos-2 and T24 cells were measured in response to dAC and photocaged dACs using Western blot analyses. The primary aim was to assess the sensitivity of DNMT1 enzyme levels to 3'- and 5'-CD. The secondary aim was to further investigate the dark activity observed in non-illuminated caged dAC treated cells.

The general observation was that illuminated 3'-CD had significant effect on DNMT1 levels. dAC and illuminated 3'-CD resulted in the disappearance of >90% of DNMT1. Interestingly, 3'-CD control treatments also had remarkable effect on DNMT1 levels. At the highest concentrations, DNMT1 levels were completely degraded, similarly to the other treatment conditions. At lower concentrations, this effect was variable between cell lines and compounds. T24 cells retained significantly higher amounts of DNMT1 compared to Saos-2 cells. Although the western blots were run a single time for each cell line and treatment condition, this conclusion is not as strong as if it was run three times. However given the similar type of response from different types of conditions (notable DNMT1 degradation) across the two cell lines, it was felt that it is not essential to repeat the western blots.

Overall, there was a significant difference between changes in DNMT1 and DNA methylation levels in response to azanucleoside treatments. The DNMT1 degrading effects of dAC and illuminated 3'-CD can occur at very low doses (as low as 0.1 μM) of the drug. However at these doses, there appears to be little or no effect on demethylation. A study by Patel and coworkers suggested that even if most of the cellular DNMT1 are degraded, residual chromatin-bound DNMT1 could still support maintenance methylation in the presence of low dose dAC.¹¹⁴ The trapping of chromatin-bound DNMT1 is more likely to occur at higher dAC concentrations and this is reflected in the corresponding greater demethylating effects.

DNMT1 levels were expected to stay unchanged in response to 'inactive' photocaged dAC in the dark. However, the Western blot analyses showed that at higher concentrations, caged dAC led to the degradation of DNMT1 even in the absence of light. This indicated that an unwanted activity took place in the dark at higher concentrations that led to the degradation of DNMT1, suggesting that photocaged dAC was no longer inactive possibly due to the release or leaking of active dAC.

3.6 Cell viability assays

3.6.1 Aims

In the final set of biological studies we aimed to evaluate the effect of photocaged dAC on cell viability. According to literature, dAC treatments at lower concentrations ($<1 \mu\text{M}$) do not affect cell viability, whereas at higher concentrations cell cytotoxicity is induced.²⁴⁴ However, the effect of photocaged dAC has not been tested on any cell lines to date. For the cell viability assays with photocaged dAC, low concentration (0.1 to $1.5 \mu\text{M}$) and high concentration (1.5 to $13.5 \mu\text{M}$) were tested.

3.6.2 Effect under low photocaged dAC concentrations

Low concentration treatments were conducted to determine any effect on cell viability of photocaged dAC analogues under biologically relevant, hypomethylating conditions that were also applied for the methylation studies and Western blot analyses.

For the assays, Saos-2, T24 and HEK293T cells were seeded on 24 well-plates in varying densities to reach around 80% confluency after 72 h. Cells were then treated with photocaged dAC (3'-CD) at three different concentrations (0.1, 0.5 and $1.5 \mu\text{M}$) and were either illuminated for 1 h at $\lambda = 365 \text{ nm}$ or kept in the dark at room temperature. dAC controls at the same concentrations and conditions were also included.

Saos-2 cell viability

Figure 3.10 depicts the cell viability curves of Saos-2 cells in response to 3'-CD and dAC at 0.1, 0.5 and $1.5 \mu\text{M}$ with or without UV illumination. Samples containing media only and media containing $1.5 \mu\text{M}$ DMSO were also included.

In general, no significant changes were observed upon treatment. A slight reduction in growth at $t = 72 \text{ h}$ corresponded to addition of dAC or 3'-CD, followed by a 1 h incubation with or without illumination at room temperature. Cells quickly recovered and reached 100% confluency within 24 h.

The uninterrupted and continuous cell growth and small error bars show that Saos-2 tolerated well UV illumination and low concentration dAC and 3'-CD (containing small volumes of DMSO) for a period of up to 24 h.

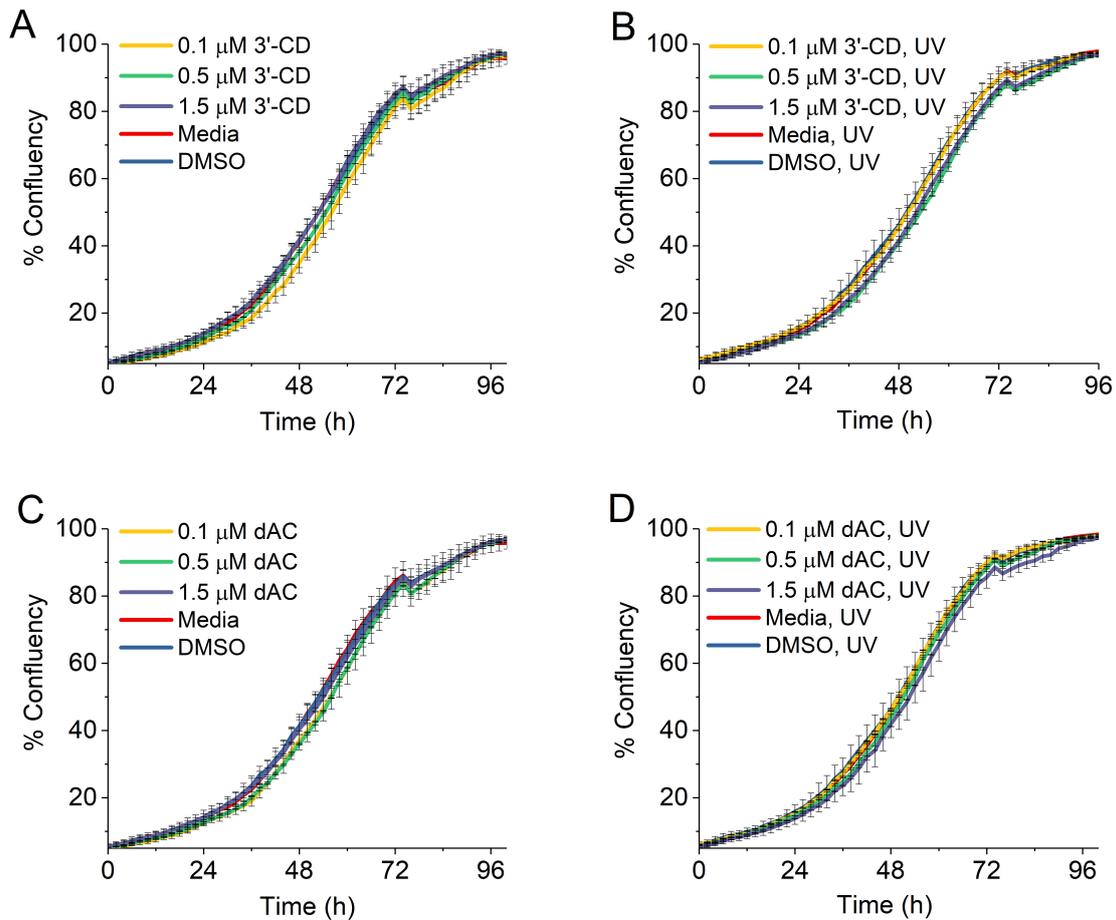


Fig. 3.10 Cell viability assays of Saos-2 cells in response to low concentration dAC and 3'-CD. (A) 3'-CD, (B) 3'-CD followed by illumination at $\lambda = 365$ nm for 1 h, (C) dAC, (D) dAC followed by illumination at $\lambda = 365$ nm for 1 h. All treatments were carried out in triplicate. The figures indicate that the growth of SaOS-2 cells were not affected by treatments with lower concentrations of dAC and 3'-CD.

T24 cell viability

The effect of dAC and 3'-CD on T24 cell viability was more noticeable (Figure 3.11). The slight drop in viability of the non-illuminated cells (corresponding to the time of the treatment at $t = 70$ h) quickly recovered in the following 24 h (Figures 3.11A and C). For illuminated cells the drop in viability was more pronounced (10%), but also recovered in the next 24 h, though with greater variability represented by the large error bars (Figures 3.11B and D). The concentration or the types of treatments did not have a significant effect on cell growth and recovery.

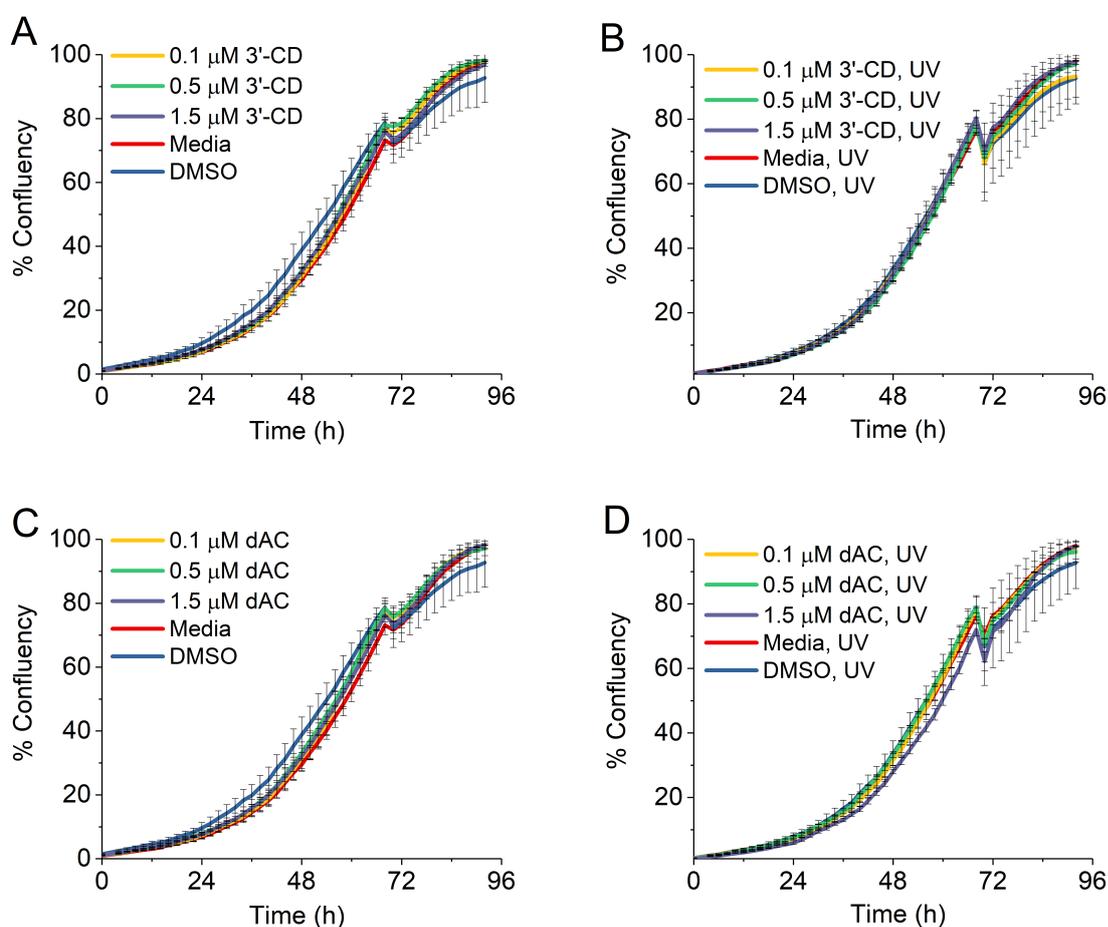


Fig. 3.11 Cell viability assays of T24 cells in response to low concentration dAC and 3'-CD. (A) 3'-CD, (B) 3'-CD followed by illumination at $\lambda = 365\text{nm}$ for 1 h, (C) dAC, (D) dAC followed by illumination at $\lambda = 365\text{ nm}$ for 1 h. All treatments were carried out in triplicate. The figures indicate that the growth of T24 cells were not affected by treatments with lower concentrations of dAC and 3'-CD.

HEK293T cell viability

Among the three cell lines, HEK293T cells had the most notable response to dAC and 3'-CD (Figure 3.12) treatments in terms of cell viability, but at the same time the least response considering methylation changes.

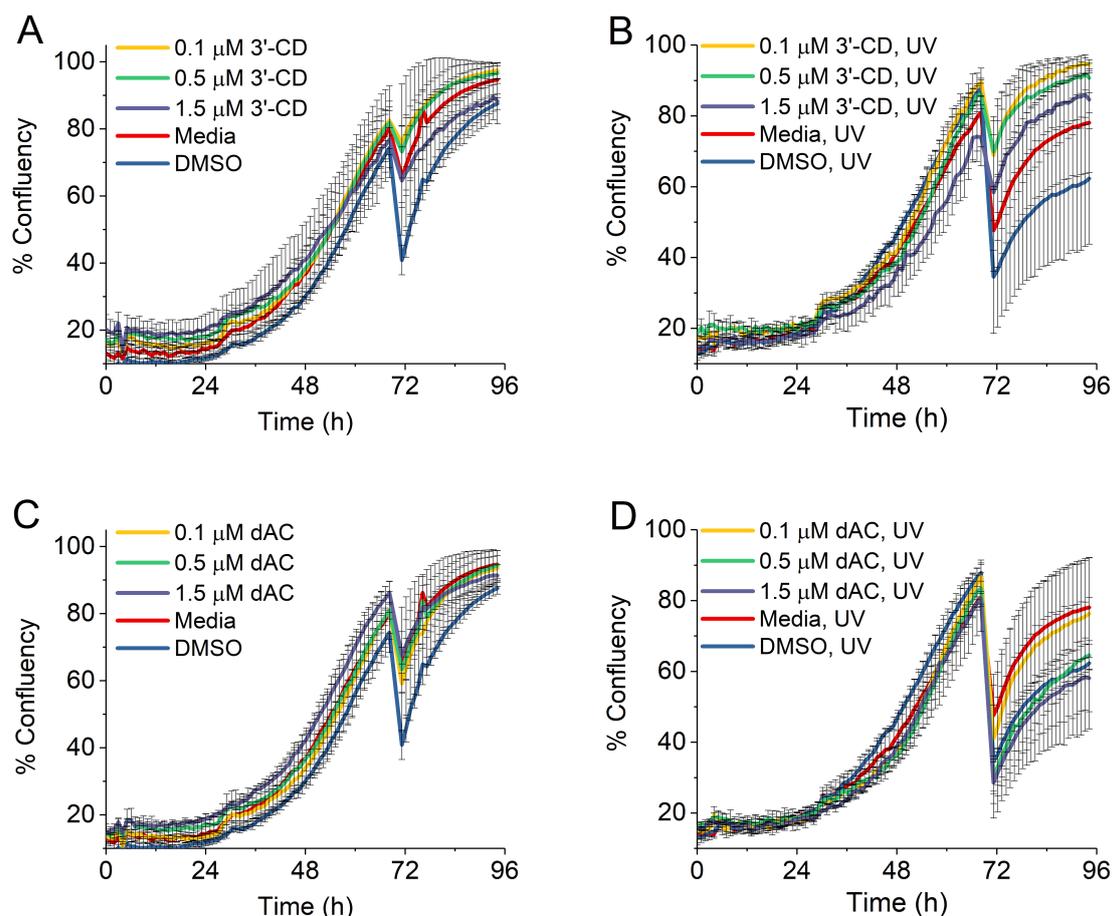


Fig. 3.12 Cell viability assays of HEK293T cells in response to low concentration dAC and 3'-CD. (A) 3'-CD, (B) 3'-CD followed by illumination at $\lambda = 365nm$ for 1 h, (C) dAC, (D) dAC followed by illumination at $\lambda = 365 nm$ for 1 h. All treatments were carried out in triplicate. HEK293T Cells were highly sensitive to dAC and 3'-CD treatments, demonstrated by the significant drop in viability after the time of the treatment ($t=70 h$). The viability slowly recovered in the next 30 hours. In the case of non-irradiated 3'-CD and dAC treatment, recovery was successful. However, irradiated samples showed a greater degree of variability possibly due to the longer term effects of treatments combined with UV radiation.

The first observation when looking at all four conditions is that treatment was followed by a drastic drop in cell viability. For non-illuminated 3'-CD, the reduction in cell

viability ranged from 7-33%, with the DMSO control having the greatest drop of all the conditions in that plate (Figure 3.12A).

DMSO and 1.5 μM CD treated cells only recovered to 80% confluency after 24 h, whereas cells treated with lower concentrations successfully recovered to nearly 100% confluency.

Given the absence of illumination in these treatment conditions, the drop in viability can be explained by the 1 h incubation of cells at room temperature, outside of the 37 °C incubator. The confluency of cells at the time of the administration may have also affected the drop in viability. DMSO-treated cells had the lowest average viability (74%) compared to cells that received 0.1 μM 3'-CD (82%) prior to administration. The recovery rate was also likely affected by the treatment condition itself. Cells that received media only recovered more rapidly compared to cells that were treated with 1.5 μM 3'-CD.

UV illumination resulted in even greater drops in cell viabilities (Figure 3.12B). The notable drop in viability is the result of the UV illumination in combination with keeping the cells outside of the 37 °C incubator for 1 h. Cells that received media only and DMSO-containing media seemed to be the most affected. One explanation could be the uneven UV distribution density among the wells as a result of the tilted positioning of the UV lamp in the UV-illuminating cabinet. Media and DMSO treated cells may have been positioned in regions in slightly higher UV density.

HEK293T cells are more sensitive to differences in UV density distribution compared to Saos-2 and T24 cells. The 3'-CD treatments had a concentration-dependent effect on cell viability, though it is difficult to conclude due to the large error bars.

dAC had similar effects on cell viability compared to 3'-CD (Figure 3.12C). However, UV illumination resulted in a significantly greater drop in viability and slower recovery rates at all concentrations (Figure 3.12D). The largest drops observed were for 0.5 and 1.5 μM dAC treated cells (53%). These cells, together with DMSO-treated cells could only recover to 60% average confluency, whereas cells that received media only and 0.1 μM dAC grew to 75% confluency after 24 h.

On the effect of DMSO on HEK293T cell growth

DMSO is an amphiphilic molecule and is a widely used solvent to facilitate the solubility of hydrophobic substances in physiological solutions. DMSO is usually well tolerated with no observable toxic effect in cells at 0.1%. At 1% or higher, toxic effects have been reported and the response will generally depend on the type of cell or cell line. DMSO treatment at higher concentrations can substantially alter the morphology and the attachment of the cells, substantially decreasing cell viability in a dose dependent manner. For example, SW620 cells showed significant change in morphology in the presence of DMSO (enlargement, elongation).²⁴⁵

The toxic effects may be in some cases due to reversible cell cycle arrest (CHO cells at 1-2% DMSO for 96 h), and in other cases it can be associated with cell differentiation to other lineages (HL-60 cells, 1-3% DMSO, 96 h) or even cell death at higher concentrations (CaCO₂ cells, >10 %, due to the release of LDH). Some other reports were suggesting that DMSO at lower concentrations had stimulating effect on cell growth. For example, at 100 μ M (0.00078% m/V) DMSO increased the proliferation of OVCAR-3, SK-OV-3 and Caov-3 ovarian cancer cell lines.²⁴⁶ Similarly, 0.05%-0.2% DMSO significantly increased the growth of RPMI-8226 myeloma cells relative to medium only control.²⁴⁷

There are no studies in literature describing any adverse effect of DMSO on HEK293T cell growth. One study looked at the effect of DMSO on the hERG channels that are expressed in HEK293 cells.²⁴⁸ The authors found that high concentration of DMSO resulted in the increase in osmolality of the bath solution which then altered the cell morphology and the biological property of the cell membrane. This would interfere with whole-cell current recordings and the authors discovered that this effect can be prevented by application of DMSO in pipette solutions. Despite that, there has been no report of adverse viability of the HEK293 cells in the study.

In our experiments, the cells treated with DMSO experienced a drop in cell viability that was even greater than the ones that contained DMSO and treatment (dAC, caged-dAC). This was intriguing as the dAC and caged-dAC treatments had the same concentration of DMSO or even lower. There could be two possible reasons for this unusual effect. One is that the concentration of DMSO in the control is higher than in the treatment solutions possibly due to uneven serial dilutions. This is highly unlikely. The second one is that the DMSO/media solution was contaminated with an unknown substance that had an effect on the cell viability resulting in a greater drop and slower

recovery. Although the DMSO control has now questionable validity, the fact that the cell growth in response to treatments in combination with DMSO and the media only control is the same suggests, that DMSO on its own does not have a significant effect on cell growth (unlike suggested by the graphs).

3.6.3 Effect under high photocaged dAC concentrations

High concentration of dAC are used less frequently due to its non-specific cytotoxic effect on cells. Given the high tolerability of Saos-2 cells towards low concentration dAC and 3'-CD, it was chosen to evaluate the effect of high concentration on cell viability.

Saos-2 cells were seeded in 24 well-plates and were allowed to grow for 24 h. They were then treated with high concentrations of dAC and 3'-CD (1.5, 4.5 and 13.5 μM) every 24 h for 3 days. Figure 3.13 shows the results of the cell viability assays.

Saos-2 cells that received media only grew to 90% confluency by the end of the viability studies. dAC had a notable effect on cell growth, however the effect did not have significant concentration dependence. All treatment concentrations resulted in a final average confluency of 40-45%. Similarly, 3'-CD also had a significant effect on cell viability, with a slight concentration-dependence - the final average confluencies reached 58%, 50% and 44% at the end of 1.5, 4.5 and 13.5 μM , respectively. The reduction in viabilities were not as large as expected from such big increases in concentration.

UV illumination did not affect the viability of cells that received media only (Media UV), but it drastically reduced viability in all other conditions. The reductions were in no particular order or concentration dependence, with final cell viabilities ranging from 22% to 35%.

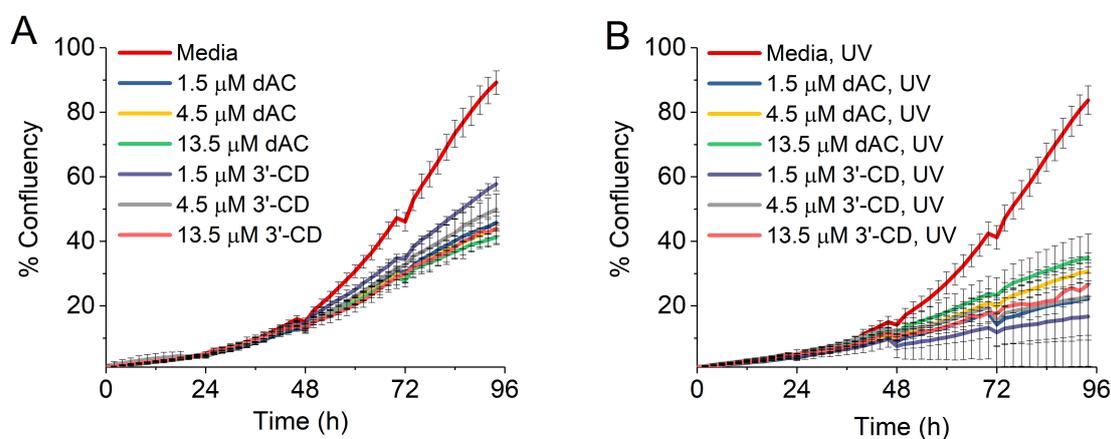


Fig. 3.13 **Cell viability assays for Saos-2 cells at high concentration dAC and 3'-CD.** Cells were treated followed by (A) dark incubation at room temperature (B) or light illumination at $\lambda = 365$ nm for 1 h at room temperature. All treatments were carried out in triplicate. The data shows that higher intensity (after 24 h of seeding, cells were treated every 24 h for 3 days) and higher concentration treatments (1.5, 4.5, 13.5 μM) have a significant effect on the growth of SaOS-2 cells, even in the absence of UV irradiation. Treatment concentrations ranging from 1.5 μM to 13.5 μM resulted in cell confluences between 40-60%. UV irradiation following treatment had an even greater effect on the cell growth, achieving cell viability between 20 and 40% after 96 h.

3.6.4 Summary of cell viability assays

In this section, the cell viabilities of various cell lines were tested in response to photocaged dAC (3'-CD) and dAC. Low concentrations were explored to determine the effect of photocaged dAC in biologically relevant hypomethylating conditions. By contrast, high concentrations were tested to establish if the photocage can block the cytotoxic effects of dAC.

At low concentration and low intensity (1 x 24 h), cell viability was not affected significantly in Saos-2 and T24 cells. By contrast, HEK293T cells experienced a considerable drop in viability due to sensitivity to temperature and UV illumination conditions during treatment.

Higher concentration and light intensity (3 x 24 h) resulted in significantly reduced viability in Saos-2 cells. The conclusion was that the photocage in caged dAC did not block the biological activity of dAC sufficiently at higher treatment concentrations, which suggested that there was a significant degree of unwanted activity in the absence of light due to the release of active dAC. This phenomenon was indicated by the methylation and western blot analyses and was then confirmed by the cell viability assays.

3.7 Preliminary cell assays on previous analogues

3.7.1 Aim of preliminary experiments

The experiments described below were carried out prior to all biological experiments discussed before. The aim of these preliminary cell experiments were to briefly assess the light-dependent activity of the N-DEACMOC-dAC **78** analogue via cell viability assays and the Infinium 450K Methylation Array. The 450K Array is a locus-specific indication of methylation changes and the analysis was carried out by Dr Sabrina Stewart.

At the beginning of the project (prior to any global methylation assay trials), a number of initial cell biological experiments were conducted on the N-DEACMOC-dAC **78** analogue, which was considered the first promising analogue amongst photocaged azanucleosides, due to its rapid uncaging rate. At the time of the cell assays, the

compound's aqueous instability was not known and fully characterised, and the 5'- and 3'-caged analogues had not been synthesised yet.

The findings from the initial cell experiments suggested the compound's instability, which then prompted further syntheses of alternative analogues. As the protocol for cell experiments changed considerably throughout the course of the project, the description of these experiments are only included for completeness.

3.7.2 Cell viability assays on N-DEACMOC-dAC

Cell viability in response to various treatment concentrations

Human osteosarcoma (SaOS2) cells were treated at 4 different concentrations of N-DEACMOC-dAC, **78**: 0.5, 1.5, 4.5 and 13.5 μM . The cells were treated every day for 24 hours followed by either a one hour illumination period at 365 nm or incubation in the dark. After the illumination or incubation in the dark, the confluence of the cells was determined every hour via an automated live cell imaging system.

Two key observations were made from the corresponding cell growth curves. First of all, the N-DEACMOC-dAC analogue on its own did not have a pronounced effect on cell growth (Figure 3.14A). However, its addition to the cells followed by illumination resulted in a concentration dependent cell killing, illustrated by the dashed lines (Figure 3.14B). As the non-treated cells were also unaffected under illumination (blue dashed line), it is confirmed that this effect was not due to UV illumination on its own.

At the time of the experiments, it was considered that N-DEACMOC-dAC treatment without illumination did not have a pronounced effect on cell growth when comparing the non-illuminated and the illuminated curves. However, after conducting further stability studies and cell experiments, the results of these cell viability curves were reconsidered. The graph on the left did suggest that a small amount of dAC was being released which resulted in a slower growth of cells at 13.5 μM treatment (brown) compared to the media control (blue).

Cell viability in response to various treatment conditions

In the same setup, the effect of media, N-DEACMOC-dAC and dAC was tested on cell viability at a single concentration of 1.5 μM (Figure 3.15). In this treatment range, the

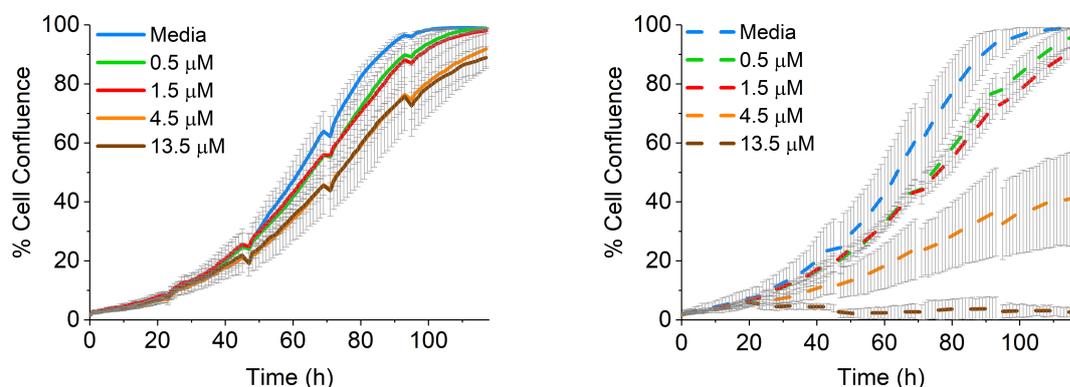


Fig. 3.14 **Cell viability of SaOS2 cells in response to N-DEACMOC-dAC** (A) with no illumination and (B) with a one-hour illumination after treatment. Caged-dAC treatment in combination with illumination resulted in a notable concentration-dependent cell killing. All treatments were carried out in triplicate

illuminated cells (red dashed line) treated by N-DEACMOC-dAC showed equivalent growth to the cells exposed to 1.5 μM dAC (black) up to about 80 hours, after which the cells have recovered.

This confirmed that the uncaging in the growth medium was successful, resulting in a similar growth to control cells up to about 80 hours. After this point, the amount of dAC released during the uncaging may not have been sufficient enough to maintain the same level of cell viability as the dAC control. At the end of this experiment, the cells were collected and their DNA was extracted in order to subject them to further analysis via 450K Methylation Array.

3.7.3 450K Methylation Array

To prove that the illuminated N-DEACMOC-dAC indeed caused reduction in DNA methylation, genomic DNA was isolated from treated (dAC, N-DEACMOC-dAC) and control cells (Media), and subjected to 450k BeadChip analysis. This technique allows the quantification of DNA methylation at over 450,000 CpG sites following bisulfite conversion. This experiment and analysis was carried out by Dr Sabrina Stewart.

The result of the microarray analysis showed that exposure to N-DEACMOC-dAC in combination with light achieved demethylation to a median of 32% 5 mC groups compared to 68% of the non-treated cells (Media) as shown by the boxplot in Figure 3.16. The value 32% is very close to reported methylation level for non-caged dAC,

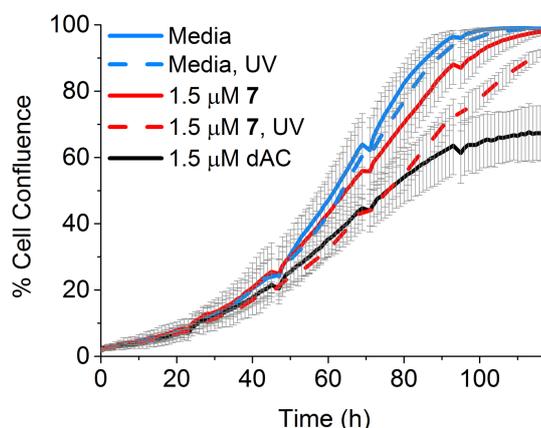


Fig. 3.15 **Cell viability of SaOS2 cells in response to various treatments conditions at 1.5 μM .** It has been shown that N-DEACMOC-dAC treatment in combination with light (red dashed line) could achieve equivalent growth to dAC control (black line) up to 80 h. Non-illuminated N-DEACMOC-dAC treatment did not have a notable effect on cell growth. All treatments were carried out in triplicate

suggesting restoration of full bioactivity after illumination and deprotection of N-DEACMOC-dAC.¹⁵⁹

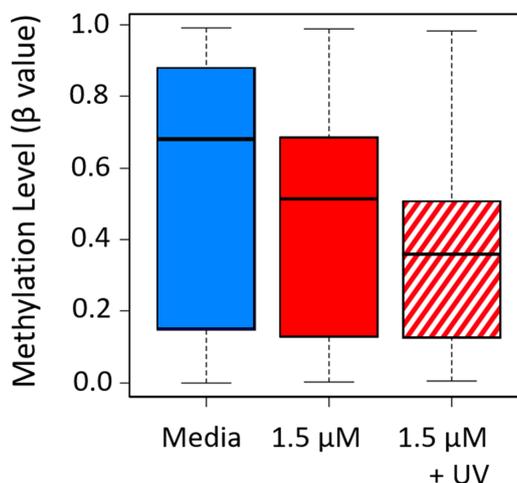


Fig. 3.16 **450K Array results.** Exposure to N-DEACMOC-dAC in combination with light resulted in demethylation to a median of 32% 5mC groups compared to 68% of the non-treated cells. However, N-DEACMOC-dAC exposure in the absence of light also resulted in notable demethylation to 52% 5 mC, suggesting that the photocage may not have been able to block activity in the dark

The microarray analysis also showed that N-DEACMOC-dAC exposure on its own also resulted in demethylation, to a median of 52%. This indicated that photocaged-dAC

on its own was not “inactive” as expected and that active dAC is potentially is being released in the dark which led to the reduction in percentage methylation.

3.8 Summary of cell biology experiments

The aim of the cell biological studies was to assess if the photocaged dAC analogues were able to achieve light-dependent demethylation. By placing a photocage on dAC, the demethylating effects of dAC were expected to be blocked. If successful, the photocaged dAC analogue could maintain methylation levels in the absence of light. Upon light illumination, the release of the active dAC should lower methylation levels in a concentration dependent manner to a similar extent as control dAC treatments.

For the methylation studies, Saos-2 cells were chosen due to their high baseline methylation level, but T24 and HEK293T cells were also included due to their known concentration-dependent sensitivity to dAC. Global DNA methylation analysis of treated cells *via* LC-MS/MS suggested that light-dependent demethylation was achieved at low concentrations ($<0.5 \mu\text{M}$). However, at higher concentration ($>0.5 \mu\text{M}$) the photocaged dAC could not maintain methylation levels, which suggested that significant activity was occurring in the absence of light.

Western blot analyses of treated Saos-2 and T24 cell extracts confirmed that photocaged dAC led to the degradation of DNMT1, the enzyme responsible for reduced methylation levels when inhibited by dAC.

Although cell viability was not significantly affected at concentrations below $1.5 \mu\text{M}$, Saos-2 cell growth was significantly reduced when photocaged dAC was administered more frequently and at higher concentrations (up to $13.5 \mu\text{M}$).

All three biological studies suggested that at the highest concentration photocaged dAC treatments, sufficient amounts of ‘active’ dAC were leaked, resulting in marked biological activity even in the absence of light. Under these conditions, decreased methylation levels were detected by global methylation analysis, reduced DNMT1 levels by western blot analyses and significant reduction in cell viability were observed in Saos-2 cells.

This is unfortunate as photocaged dAC analogues should remain stable and not release dAC in the absence of light. The source of instability could be the chemical or enzymatic hydrolysis of the caged compounds in media and the cellular environment. In order to

reassess the stability of the photocaged compounds in such conditions, further studies were carried out using HPLC and mass spectrometry analyses.

3.8.1 Limitations and suggestions for further improvements

The setup of the current cell experiments had a number of limitations that could be addressed in the future cell experiments.

First of all, the UV-irradiation system used in these experiments was a benchtop UV Cabinet that contained the UV lamp on top with a slanted position. This resulted in an uneven distribution of light density that reached the wells. Since the wells were 20 cm from the lamp and the measured variation in density was around 5-10 $\mu\text{W}/\text{cm}^2$, we went ahead with this system as at the time a horizontal UV cabinet was not available. It was confirmed that the variation in UV power density did not have an effect on cell growth when the cells received no treatment.

For these set of experiments, the treatments of the wells were in a non-randomised format (see Experimental section) ie. the similar treatment concentrations were adjacent to each other in order to speed up the addition of treatment to the wells. However this setup could potentially skew the results – the wells with higher power density (closer to the lamp) could have a faster uncaging compared to wells that receive lower power density from the lamp. In order to avoid this bias, future cell treatments should include randomisation in the position of the treatment wells.

Chapter 4

Stability of Photocaged Azanucleosides

4.1 Introduction

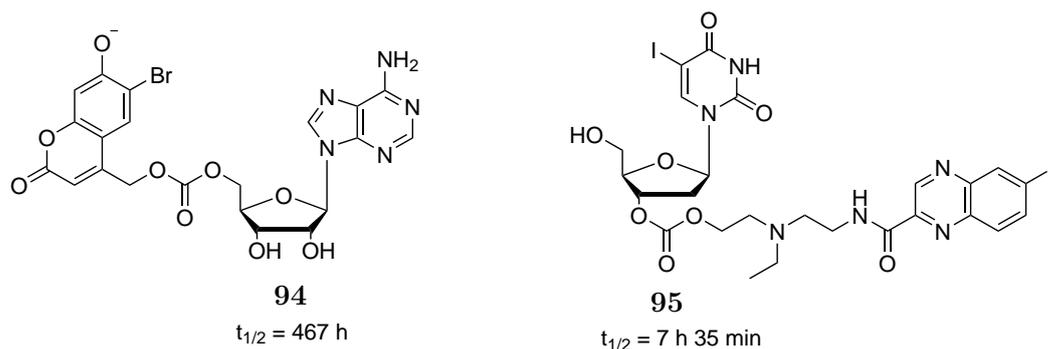
Photocaged azanucleosides are biologically inactive forms of azanucleosides. The stability of a photocaged azanucleoside is crucial in maintaining its inactivity and light tunable nature. The biological studies described in Chapter 3 suggested that when cells were treated with photocaged dAC, active dAC was being released due to hydrolysis of the carbonate linker resulting in unwanted biological effects.

On the contrary, our previous HPLC stability studies (Chapter 2) suggested that the 3'- and 5'-caged analogues did not release active dAC. Incubation of these analogues did not give rise to detectable dAC peaks, suggesting that the carbonate link between dAC and the photocage remained intact. However, over 24 h of incubation at 37 °C, the main caged dAC peak degraded into additional smaller peaks with similar retention times, which was indicative of the triazine ring opening.¹⁵⁵

The main aim of this chapter was to further investigate the stability profiles of photocaged dAC analogues. The first objective was to detect the source and mechanism of instability that was observed in the cell biological studies. The second objective was to quantitatively characterise the degradation profiles to determine the exact amount of any active dAC that was being released.

4.2 Literature on carbonate link stability

Prior to describing experimental analysis, a literature survey on the stability of the carbonate linkage is provided. The stabilities are variable in different literature reports. For example, a 5'-coumarin-caged adenosine analogue **94** with a carbonate linker synthesised by Suzuki *et al.* had a half-life of 467 h (over 19 days).¹⁹¹ On the contrary, André *et al.* found that carbonate linkers have moderate stabilities that are poorer than carbamate but stronger than ester bonds.²²⁰ When exploring various linkers between 5-iodo-2'-deoxyuridine and a melanoma tracer, the authors observed that the carbonate-linked analogue **95** had a half-life of 7 h 35 min in phosphate-buffered saline (PBS) buffer at 37 °C. In cell culture medium and human serum the half-lives were even lower with 6 h and 4 h, respectively. The hydrolysis of carbonate bonds in serum could be accelerated by enzyme in the serum such as carboxyl esterases and plasmin enzymes.²⁴⁹



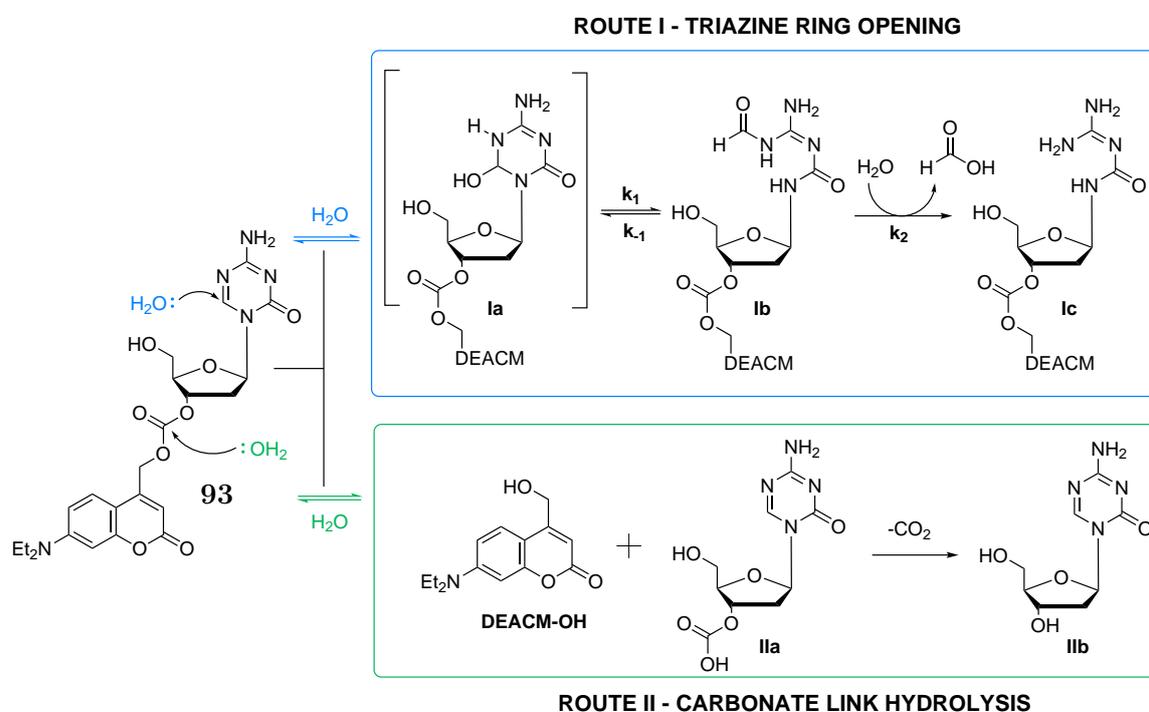
Scheme 4.1 Reported stabilities of carbonate linked nucleosides in literature^{191,220}

The stability of carbonate linkers greatly depends on the structures of the linked analogues and the incubation environment. To date, no study has been reported on the stability of carbonate-linked dAC analogues. Despite the known instability of dAC caused by the triazine ring opening, our aim was to determine the stability of the carbonate linker in photocaged dAC analogues.

4.3 Hypothesis on the sources of instability and aims

Based on the literature studies, we hypothesised that two sources of instability could account for the degradation of photocaged dAC analogues (Scheme 4.2). The first source is the hydrolytic opening of the triazine ring (Route I) known in the literature.^{154,155} In this proposed mechanism, water attacks at the electrophilic C6 position of the triazine ring of compound **93**, forming intermediate **Ia**. This intermediate is in equilibrium with analogue **Ib** where the triazine ring is partially open. Irreversible ring opening occurs upon attack of another water molecule, releasing formic acid and giving analogue **Ic**.

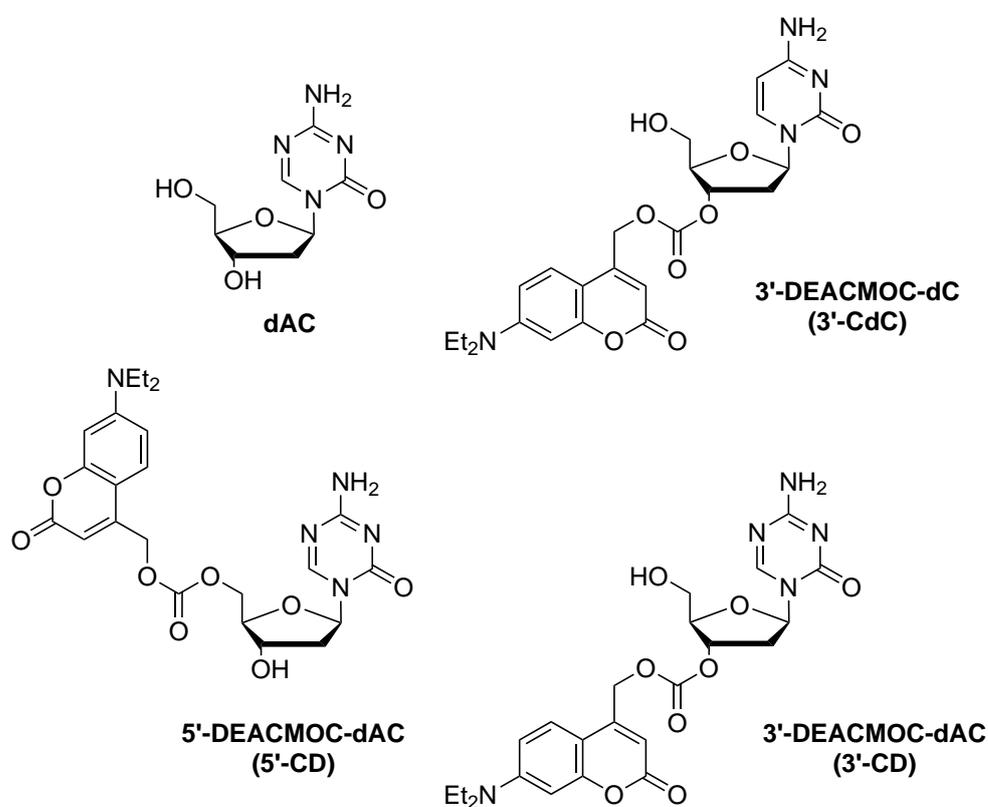
The second possible source of instability is the hydrolysis of the carbonate link (Route II) that could be initiated chemically or enzymatically. The hydrolysis may release the free DEACM-OH phototag first, leaving a 3'-carboxylated dAC analogue (**IIa**) that eventually decarboxylates to give active dAC (**IIb**). The rate of this hydrolysis depends on the chemical nature of the two linked components, the temperature and pH of the buffer.^{220,250}



Scheme 4.2 Possible mechanisms of instabilities in photocaged dAC

The aim was to test the hypothesis by quantifying the rates of triazine ring opening and carbonate link hydrolysis. To determine rates for triazine ring opening, the non-photocaged dAC was selected. By comparison, for the rate of carbonate link hydrolysis a reference compound was chosen that is related in structure to 3'-DEACMOC-dAC (3'-CD) but does not have triazine ring instability which would complicate analysis. The reference compound is 3'-DEACMOC-dC (3'-CdC) which contains the hydrolytically stable cytosine base instead of the triazine ring (Scheme 4.3)

3'-DEACMOC-dC (3'-CdC) was synthesised by Master's student Gurpreet Viridi. As additional reference, 5'-DEACMOC-dAC (5'-CD) was also selected (Scheme 4.3)



Scheme 4.3 Structures of compounds used for stability studies

4.4 Stability studies *via* HPLC

The stabilities of dAC and 3'-CD at 37 °C in water were measured and quantified *via* HPLC (Chapter 2). Two additional concentrations of 3'-CD were also tested. The 24 h stability profiles are shown in Figure 4.1. The incubation of 100 μM dAC at 37 °C led to the degradation of 55% of the compound after 24 h (Figure 4.1A). The measured half-life was 11 h, which is in close agreement with reported literature values ($t_{1/2} = 10$ h).¹⁵⁵

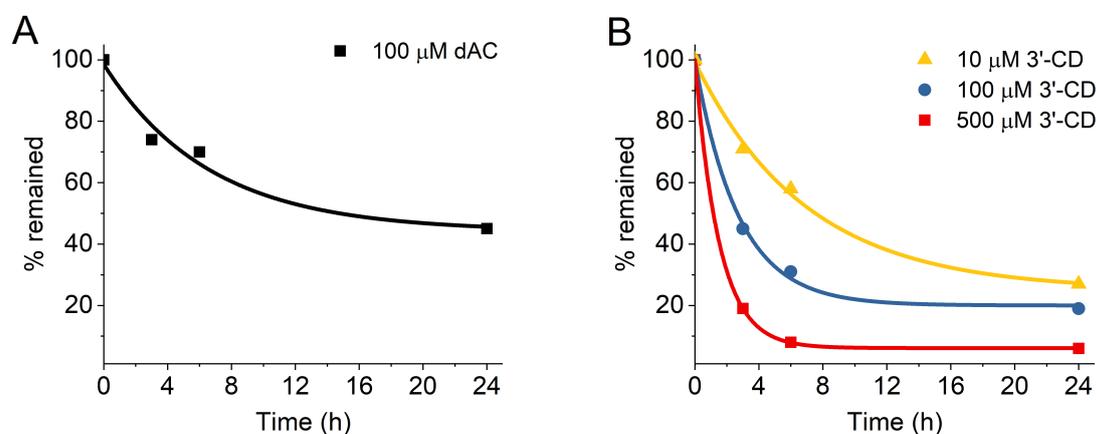


Fig. 4.1 HPLC stability analysis of dAC (A) and photocaged dAC at different concentration (B) at 37 °C

The analysis of 3'-CD stability was carried out at three different concentrations (10, 100 and 500 μM) to determine if the concentration affects degradation rates (Figure 4.1B). The results showed that increasing the concentration of 3'-CD significantly increased degradation rates. The corresponding half-lives of 10, 100 and 500 μM 3'-CD stability profiles were $t_{1/2} = 5.5$, 3.5 and 1.5 h respectively. One possible explanation for these surprising results is the aggregation or limited solubility of 3'-CD in aqueous solutions at higher concentrations.

The concentrations used in HPLC stability studies were higher than those used in cell experiments (0.1 – 1.5 μM). As HPLC has a low detection limit of 10 μM with a likely negative impact on accurate measurement of stability profiles, a decision was made to switch to alternative analytical methods that allowed the detection of submicromolar concentrations with high accuracy. LC-MS/MS was identified as the method of choice.

4.5 Stability studies *via* mass spectrometry

4.5.1 Triple quadrupole mass spectrometer

LC-MS/MS is a highly sensitive and accurate method to detect and quantify small molecules. The technique also provides additional information of molecular weight and enables the separation of compounds based on their unique fragmentation profile.

For the stability studies, an LC system coupled to a triple quadrupole mass spectrometer at UCL School of Pharmacy was used. Figure 4.2 shows a schematic diagram of a triple quadrupole mass spectrometer.²⁵¹

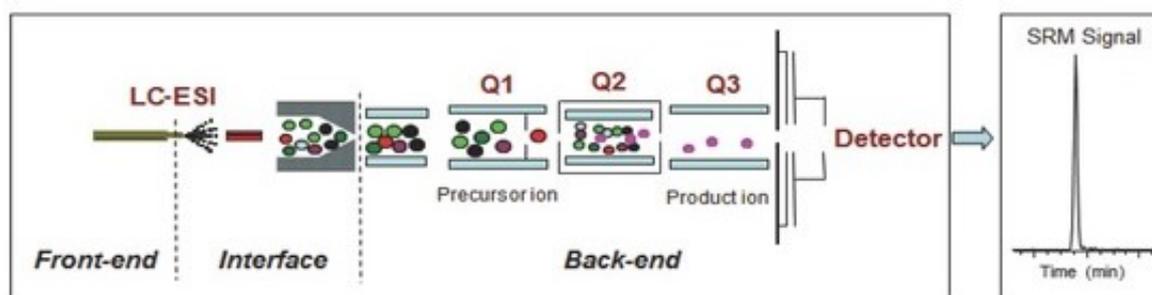


Fig. 4.2 Schematic diagram of a triple quadrupole mass spectrometer²⁵¹

During LC-MS/MS analysis, a sample mixture is separated on an LC column followed by ionisation of the individual analytes. Mass detection is carried out in three steps, *via* sequential filtering through three chambers that contain the quadrupole mass filters. Quadrupoles consist of four cylindrical rods set parallel to each other that are responsible for the filtering of sample ions based on their mass to charge ratio (m/z). The first quadrupole selects the ions of desired m/z and transmits them to the second quadrupole. There, the ions are further fragmented, and the third quadrupole scans the entire m/z range to determine the sizes of the fragmentation products that subsequently reach the detector.

A triple quadrupole mass spectrometer features various scan modes. For the stability studies, multiple reaction monitoring (MRM) was used (Experimental Section, page 214) to selectively detect and quantify dAC, photocaged dAC and phototag fragment ions in the stability mixture at high sensitivity.

4.5.2 MS stability studies in water

Stability studies at room temperature

In the first round of stability studies, the selected analogues (1.5 μM dAC, 3'-CD, 5'-CD and 3'-CdC) were incubated in water at room temperature and their stabilities were monitored for 6 h (Figure 4.3). The degradation of dAC showed an unusual pattern (Figure 4.3A). After 30 min of incubation at room temperature, only 58% of the total dAC was detected. This was followed by a significantly slower decay, reaching 44% of the starting dAC concentration at the end of the 6 h incubation period. One explanation could be the equilibrium between the dAC ($M_W = 228$) and its partially ring opened form ($M_W = 246$) (Scheme 4.2). The exchange between the two species is potentially slow enough to allow their separate detection via MS. The slow decay from 58% to 44% of dAC potentially corresponded to the irreversible opening of the triazine ring.

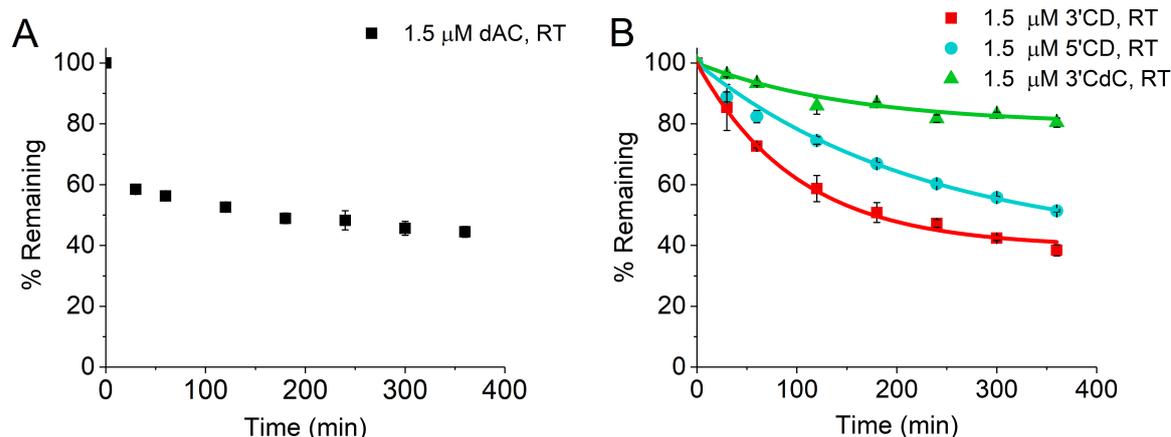


Fig. 4.3 Stability profiles of dAC and caged analogues in water at RT measured by MS

The stabilities of 3'-CD, 5'-CD and 3'-CdC at room temperature are shown in Figure 4.3B. 3'-CdC was the most stable, with 80% present after 6 h of incubation ($t_{1/2} = 16.5$ h). The moderate disappearance of 3'-CdC was due to the carbonate bond hydrolysis. By comparison, the rate of decay for 3'-CD was faster ($t_{1/2} = 180$ min), leaving 38% of the intact compound at the end of the 6 h incubation period. The faster decay was likely due to the combination of carbonate bond hydrolysis and triazine ring opening. The stability of 5'-CD was surprisingly better than that of 3'-CD with 51% still intact compared to 38% of 3'-CD.

Stability studies at 37 °C

The next round of stability studies was carried out at 37 °C to better mimic cell biological conditions. The measured stability of dAC at 37 °C was markedly different from the profile obtained from the room temperature (Figure 4.4). After 6 h of incubation, 60% of the intact dAC was detected. The degradation was fitted to an exponential decay curve with a half-life of 11.5 h, which was in close agreement with our previous HPLC stability studies and reported literature values.¹⁵⁵ We postulated that the exchange between dAC and its partially ring-opened form at 37 °C was more rapid and that the two species were indistinguishable using MS analyses.

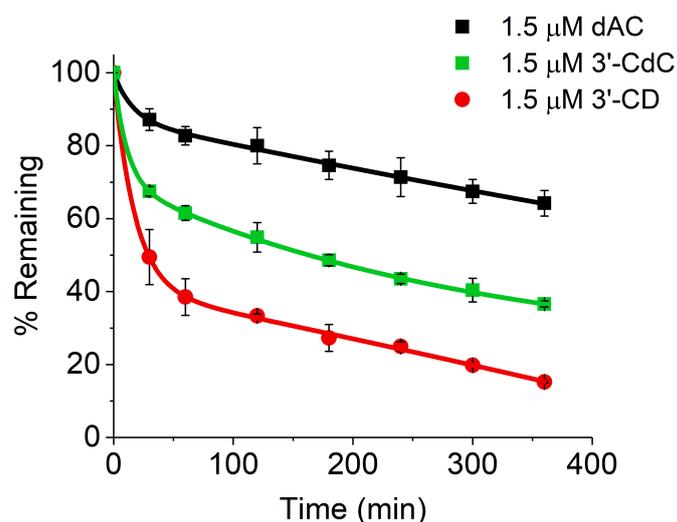


Fig. 4.4 MS Stability profiles of dAC, 3'-CD and 3'-CdC in water at 37 °C

3'-CdC was less stable at 37 °C than at room temperature. Considerable amounts of the analogue disappeared over 6 h, leaving 37% of intact 3'-CdC ($t_{1/2} = 180$ min), indicative of a faster rate of carbonate hydrolysis at higher temperature. The degree of decay of 3'-CD was also higher than dAC and 3'-CdC at 37 °C and 3'-CD at room temperature. Indeed, after 6 h, only 15% of the intact 3'-CD was detected (Figure 4.4), with a measured half-life of 30 min. The fast dissociation of 3'-CD at 37 °C is likely due to the combination triazine ring opening and the carbonate link hydrolysis, which is accelerated by the increased kinetic energy resulting from the rise in incubation temperature (from room temperature to 37 °C).

Determination of dAC released during 3'-CD incubation at 37 °C

The degradation of 3'-CD at 37 °C was further investigated in order to determine the amount of dAC released. Multiple reaction monitoring (MRM) enabled the quantification of the three main analytes (dAC, 3'-CD and DEACM-OH phototag) at each time points of the stability analysis as shown in Figure 4.5.

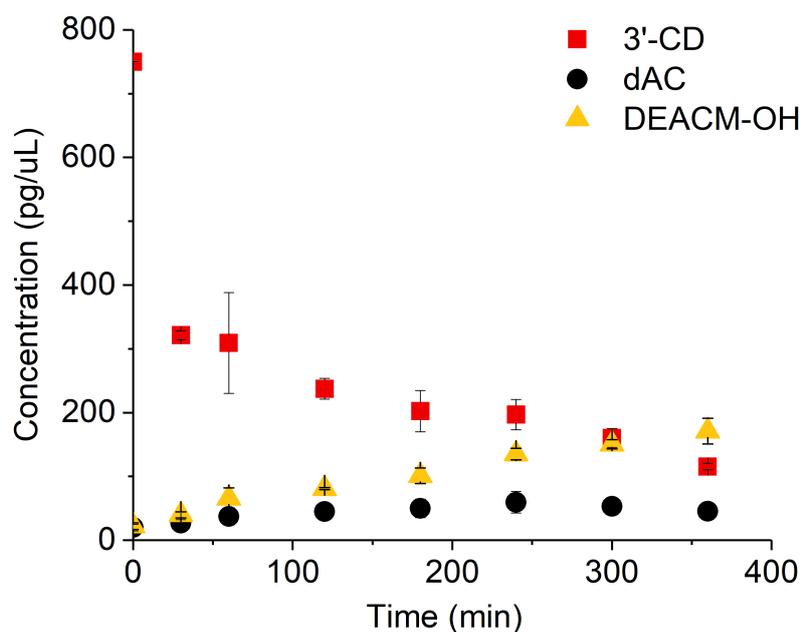
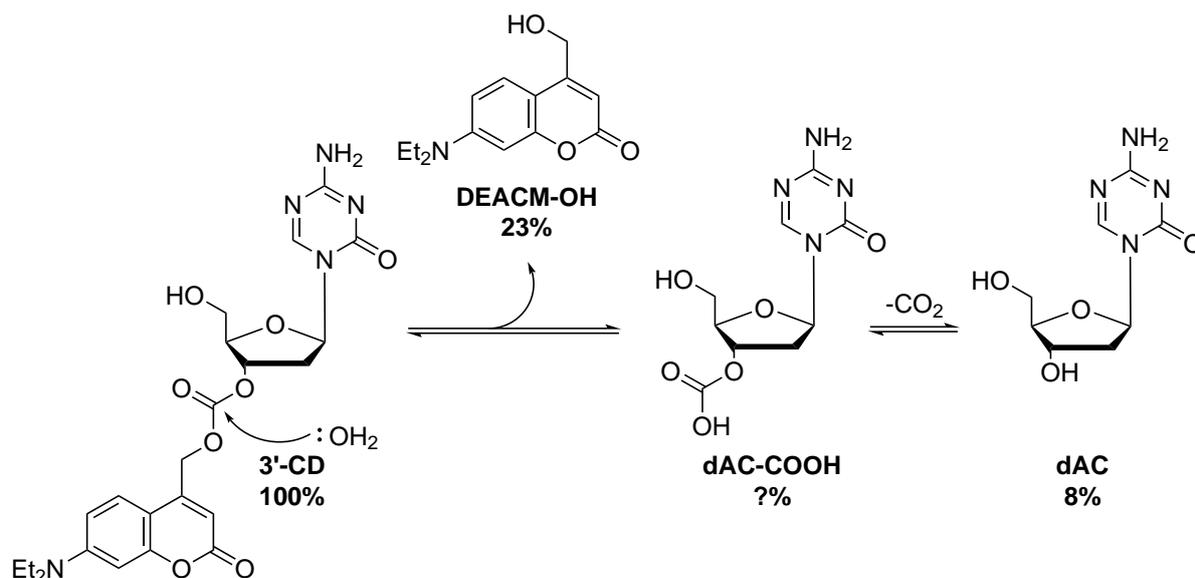


Fig. 4.5 Quantification of analytes during 3'-CD incubation in water at 37 °C

Strikingly, the rapid disappearance of 3'-CD did not correspond to an equally rapid and matching appearance of dAC and the DEACM-OH phototag. The initial 3'-CD concentration of 750 pg/μL (1.5 μM) decreased to 115 pg/μL after 6 h of incubation (15% of the initial 3'-CD). On the contrary, the concentration of DEACM-OH phototag gradually increased during the course of the stability study, reaching a final concentration of 171 pg/μL after 6 h. This suggests that only 171 pg/μL of the original 3'-CD (23%) underwent the first steps of direct carbonate bond hydrolysis that released the DEACM-OH phototag as shown in Scheme 4.4.

Considerably less of the 3'-CD was directly converted to active dAC. The concentration of dAC increased slower than DEACM-OH, reaching a maximum amount of 60 pg/ μ L at $t = 240$ min followed by a gradual decrease, which could be due to the triazine ring opening. These measurements suggest that approximately 8% of the photocaged dAC is directly converted to active dAC as a result of chemical hydrolysis during incubation at 37 °C. A mechanistic summary of this process is shown in scheme 4.4.



Scheme 4.4 Mechanism of carbonate link hydrolysis

The rate of carbonate link hydrolysis greatly depends on the pH and components of the buffer.²²⁰ The stability studies up to this point were carried out in HPLC grade water for simplicity of detection. However, the pH of HPLC grade water (pH=8.4) was higher than the pH of the physiological environment (pH=7.4) or the pH of the cell culture media (pH=6.9-7.3) used in the cell biological studies.

4.5.3 MS stability studies in HEPES buffer pH 7.2

To mimic the pH conditions of the cell biological studies, the next stability experiments were conducted in HEPES buffer at pH 7.2. HEPES is widely used in cell culture mainly because it maintains physiological pH well, despite changes in CO₂ concentrations. The reduced complexity of HEPES buffers compared to PBS and other cell culture media also enables its direct use in mass spectrometry analyses.

As a drawback, the presence of buffer made the detection of dAC using MS more challenging. In particular, the buffer and dAC likely eluted at similar retention times, leading to the interference of buffer molecules with the detection of dAC analytes. Nevertheless, the presence of HEPES buffer did not interfere with the detection of other photocaged dAC analogues and DEACM-OH phototag. Despite the limitation in the detection of dAC analytes, the photocaged analogues were subjected to stability studies in HEPES buffer. Figure 4.6 shows the stability profiles of the analogues at room temperature (A) and at 37 °C (B).

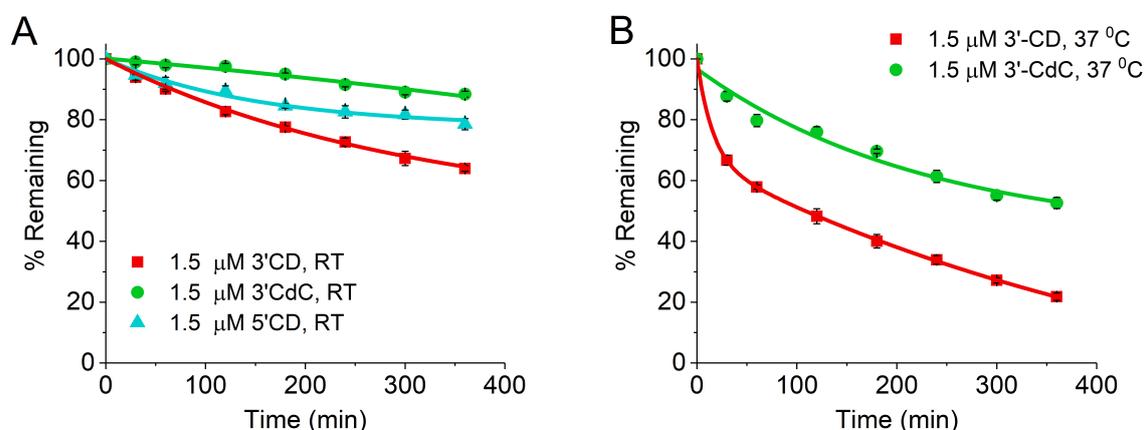


Fig. 4.6 MS Stability profiles of photocaged analogues in HEPES pH 7.2 at (A) room temperature and (B) 37 °C

In buffer, the stabilities of photocaged analogues (3'-CD and 5'-CD) at room temperature were significantly improved (Figure 4.6A). 64% of 3'-CD and 78% of 5'-CD were intact after 6 h of incubation. In the case of 3'-CdC, the compound also showed higher stability in buffer. This suggests that HEPES buffer - which had a lower pH than water - slowed down hydrolysis of the carbonate linkage, in line with literature observations.²²⁰

HEPES also increased the stability of compounds at 37 °C when compared to water (Figure 4.6B). The disappearance rate of 3'-CD in HEPES was higher compared to the room temperature studies, however they were slightly improved compared to the rates observed in water at 37 °C. 3'-CdC was still considerably more stable than 3'-CD. After 6 h of incubation, 50% of the initial 3'-CdC was still intact.

4.6 Summary of stability studies

The aim of the stability studies was to characterise the mechanisms of instability of photocaged dAC analogues and to quantify their degradation products. Mass spectrometry analysis was chosen as the analytical method of choice due to its accuracy and higher sensitivity of detection compared to HPLC.

The stabilities of photocaged dAC analogues were tested in water and HEPES buffer pH 7.2 at room temperature and 37°C. In addition, the stabilities of dAC and a 3'-caged dC analogue were also measured to separately determine the rate of triazine ring opening and carbonate link hydrolysis, respectively. Table 4.1 summarises the measured half-lives from the MS results.

Solvent	Temperature	dAC	3'-CdC	3'-CD	5'-CD
Water pH 8.4	RT	30 min	12 h 50 min	3 h	6 h
	37°C	11 h 30 min	3 h	30 min	N/A
HEPES pH 7.2	RT	N/A	38 h 30 min	11 h 30 min	16 h 30 min
	37°C	N/A	5 h 50 min	1 h 40 m	N/A

Table 4.1 Summary of half-lives from MS stability studies

The stability data provided valuable information on the two main mechanisms of fragmentation that contributed to the degradation of 3'-CD.

Triazine ring-opening rates were extracted from the stability profiles of dAC. The measured half-life of dAC degradation in water ($t_{1/2} = 11.5\text{h}$) was in agreement with reported literature values.¹⁵⁵ Incubation at 37°C led to the degradation of 36% of the initial amount of dAC. This suggests that in photocaged dAC, this portion of the observed decay was due to triazine ring opening.

The rates of hydrolysis for the carbonate linkage were determined from the examination of the 3'-CdC stability. The structure of 3'-CdC is identical to that of 3'-CD except for the stable cytosine ring instead of a triazine ring that undergoes ring opening. The measured half-life of 3'-CdC degradation was 5 h 50 min in buffer and 3 h in water, which suggests that carbonate link hydrolysis rates were greater than triazine ring opening rates. After 6 h of incubation at 37°C, 48% of the carbonate-linked analogue was degraded in buffer.

The measured rates of decay for dAC and 3'-CdC confirmed our initial hypothesis that the overall rate of instability in 3'-CD is the combination of the rates for triazine ring opening and carbonate link hydrolysis, respectively. Indeed, the calculated sum of the two degradation curves of dAC and 3'-CdC ('predicted' 3'-CD stability) was close to the experimentally measured 3'-CD degradation curve (Figure 4.7).

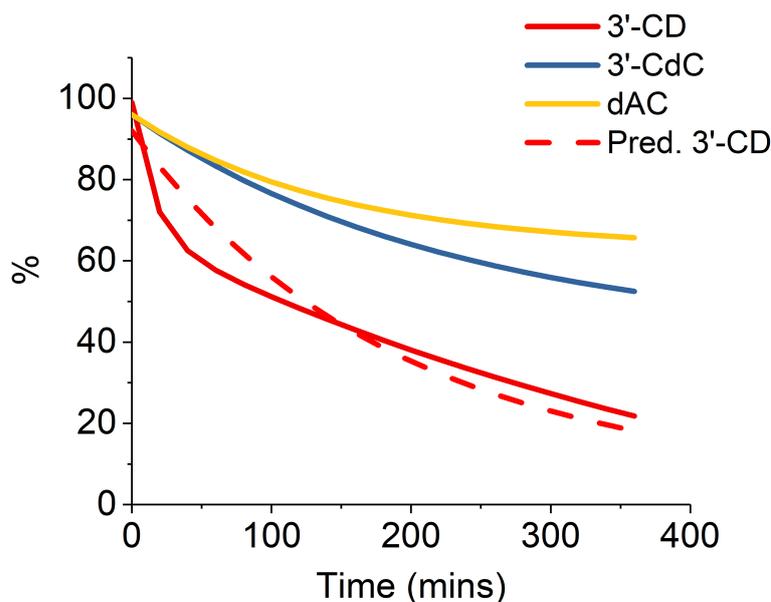
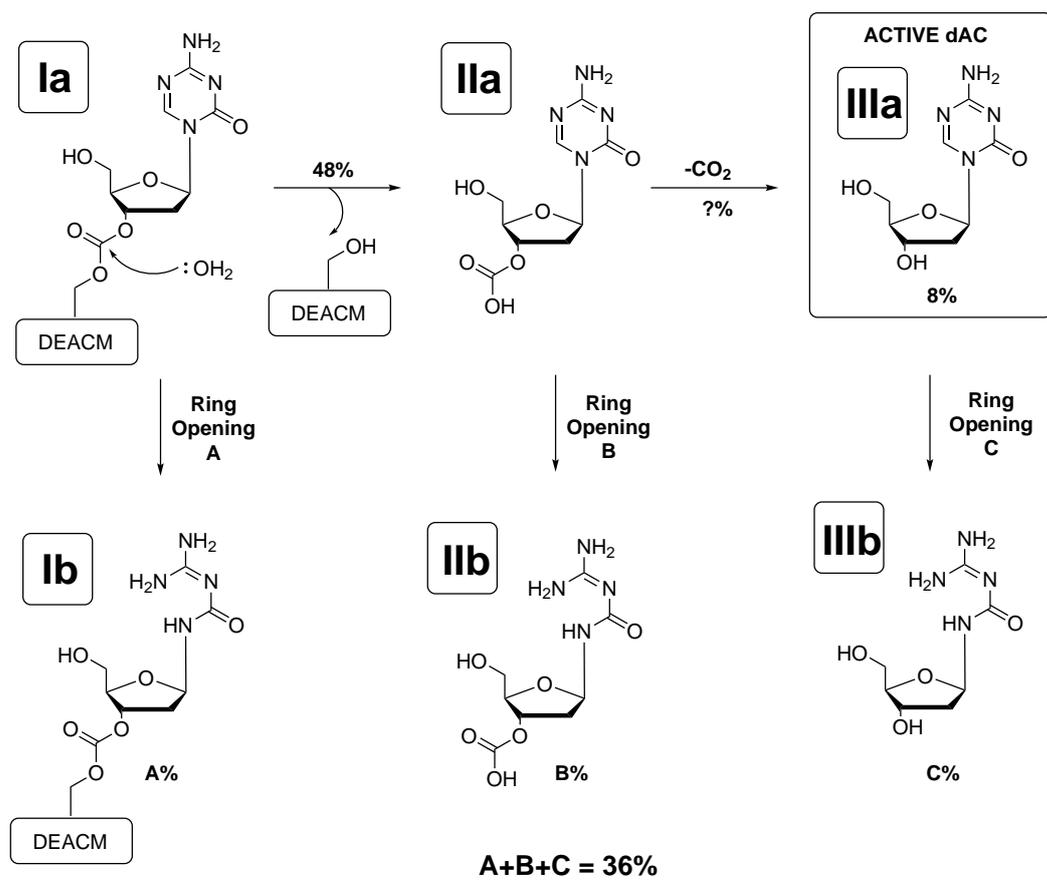


Fig. 4.7 Comparison of calculated and experimentally measured degradation curves of 3'-CD

Further analysis of the 3'-CD stability profiles provided additional insights on the degradation mechanisms. Despite the rapid rate of disappearance of the photocaged analogue 3'-CD (15% remaining after 6 h), only 8% of the active dAC was released during a 6 h incubation period. This observation may be accounted for by the degradation mechanism that also features relative rates of decays for the different pathways (Scheme 4.5).

In the proposed scheme, 3'-CD (**1a**) can either undergo 'Ring Opening A' (to give **Ib** to an unknown extent - A%) or carbonate link hydrolysis (maximum of 48%) to give the 3'-carboxylated dAC analogue (**IIa**). Decarboxylation of this analogue is thought to proceed slowly and can directly release active dAC (**IIIa**). However, 'Ring Opening B' of the carboxylated analogue (to give **Ib**) can also compete with and prevent the release of active dAC.



Scheme 4.5 Proposed mechanism of degradation of photocaged dAC

In addition, after the release of the active dAC, it can also be further degraded by 'Ring Opening C' to give **IIIb**. The combined effects of the three ring opening steps (A/B/C) contribute to the degradation of 36% photocaged dAC *via* ring opening. The slow decarboxylation rates and the competing ring opening mechanisms at each step of the carbonate link hydrolysis could be an explanation for the limited amount of active dAC released.

In summary, a viable interpretation of the cell biological studies was proposed based on the results obtained from the stability studies. The unwanted activity was due to the release of >8% active dAC at higher caged dAC treatments that resulted in comparable effects at equivalent dAC concentrations.

Chapter 5

Non-hydrolysable Azanucleosides

5.1 Introduction

The previous chapters were focused on the synthesis, photochemical characterisation and stability assessment of photocaged analogues. The aim of this section was to synthesise and evaluate the bioactivities of non-hydrolysable and non-photocaged forms of azanucleosides.

It has been long established that dAC is a DNMT1, and in order to exert its bioactivity, it first has to be incorporated into the replicating DNA.²⁵² However, a study published by Ghoshal and coworkers suggested that this might not be the case.¹¹⁷ In their work, the authors demonstrated that dAC could target DNMT1 via a proteosomal degradation pathway even in the absence of DNA replication, suggesting that dAC may exert bioactivity even when it is not incorporated into the DNA.

Taking this finding into account, the following questions arose considering photocaged azanucleosides: *If the photocage (on the 5'- or 3'-position) blocks the incorporation of dAC into DNA, but the „free flowing” dAC could still be active (ie. degrade DNMT1) in the absence of DNA replication, does it mean that photocaged-dAC can also exert bioactivity in its „free flowing” form?*

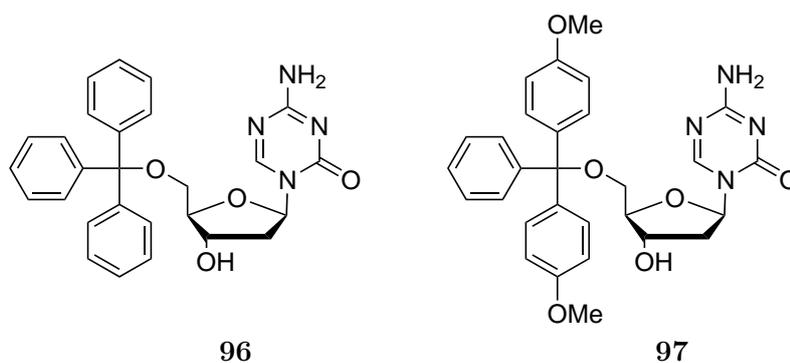
The triazine ring of dAC is thought to contribute towards the DNMT1 blocking activity of dAC. Since the photocage (that is conjugated either on the 3'- or 5'- position) is not sterically blocking the triazine ring of dAC, the suggested theory above could be a plausible mechanism.

Due to the problematic aqueous stability profiles of the photocaged azanucleosides, it is challenging to determine, whether the observed bioactivity (ie. demethylating effect) in the absence of illumination was solely due to the release of dissociated dAC as a result of carbonate bond hydrolysis, or also due to the activity of the „free flowing”, intact photocaged analogue.

To further investigate, it was envisioned that a non-hydrolysable form of dAC would provide further insights to the question above. A dAC analogue with a non-hydrolysable linker would be able to distinguish between replication dependent (ie. bioactivity that requires dAC incorporation into the DNA) and independent activities (ie. activity of „free flowing” analogues) of dAC. Synthesis of Non-hydrolysable forms

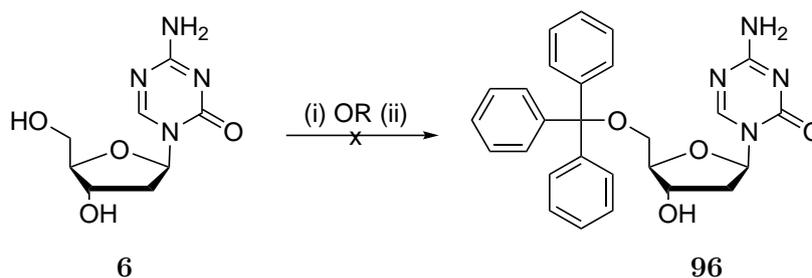
5.2 Synthesis of non-hydrolysable Azanucleosides

Example structures of a non-hydrolysable forms of dAC include tritylated dAC (**96**) and DMT-conjugated dAC (**97**) (Scheme 5.1). Ether linkers are known to be significantly more resistant to aqueous hydrolysis compared to carbonate bonds..²²⁰ Trityl ethers also show good stability in mild acidic or basic media which make them good candidates to be used in total syntheses of related targets.²⁵³ In this section, the synthesis of both analogues was attempted.



Scheme 5.1 Structures of non-hydrolysable Azanucleosides

The synthesis of tritylated dAC (**96**) was attempted in several rounds. First, a protocol described by McGuigan and coworkers was followed (Scheme 5.2, (i)).²⁵⁴ in which dAC was treated with 2.2 equivalents of trityl chloride in the presence of DMAP and pyridine at 80 °C. Unfortunately the reaction did not yield any products. The same protocol was repeated at room temperature, but again no product spot was detected.

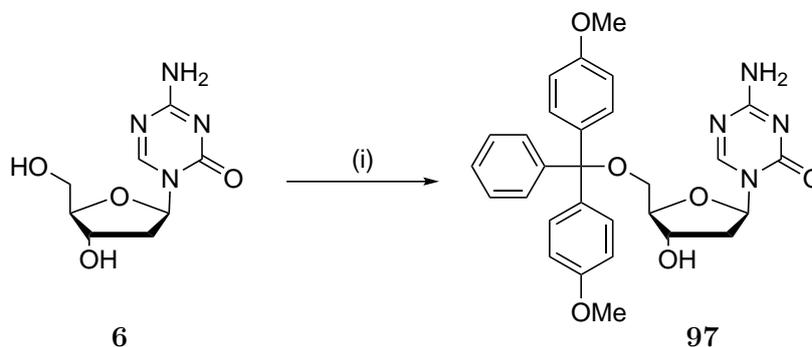


Scheme 5.2 Synthesis of Trityl-dAC **96**. *Reagents and Conditions:* (i) 2.2 eq TrCl, DMAP, pyridine, 80 °C, 24 h; (ii) 1.1 eq TrCl, DMAP, imidazole, DMF, rt, 24 h

As a control, AC was used instead of dAC as a starting material (as described in the original protocol), but this time a clear product spot appeared on the TLC, which was then later confirmed by LC/MS to be the product spot. In summary, this method was only suitable for AC tritylation, but not dAC tritylation.

An alternative method described by Andre and coworkers was trialed, in which TIPS protection of 5'-OH of nucleosides was conducted (Scheme 5.2, (ii)).²²⁰ In this method, dAC was treated with 1.1 equivalents of trityl chloride (instead of TIPS-Cl as described in the original protocol) in the presence of DMAP, imidazole and DMF at room temperature. Unfortunately this reaction did not yield any product either.

Given that the tritylation of dAC was unsuccessful, the conjugation of an alternative group, DMT was attempted next (Scheme 5.3). A protocol described by Kim and coworkers involved the treatment of dAC with 1.1 equivalents of DMT-chloride in the presence of pyridine at room temperature, followed by 80 °C.²⁵⁵ The reaction successfully gave DMT-dAC in 52% yield.



Scheme 5.3 Synthesis of DMT-dAC **96**. *Reagents and Conditions:* (i) 1.1 eq DMT-Cl, pyridine, 80 °C, 24 h

5.3 Stability analysis of non-hydrolysable Azanucleosides

The stability of the DMT-dAC at 37 °C in water was assessed using HPLC analysis and is shown in Figure 5.1. In the HPLC trace, two main peaks can be detected, one of them corresponds to DMT-dAC ($R_t=17$ min) and the other one to dAC ($R_t=5$ min), which is likely an impurity from the synthetic mixture.

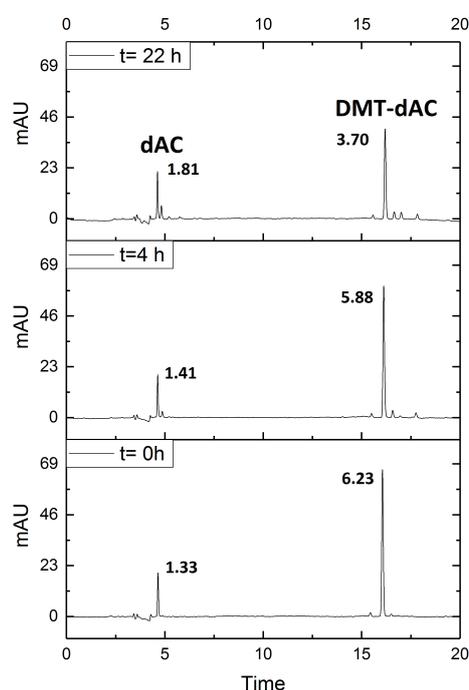


Fig. 5.1 HPLC Stability analysis of DMT-dAC, **97**.

During the incubation, the dAC peak did not increase significantly after 4 and 22 hours, according to the peak integral values. This suggests that there was no rapid dissociation from DMT-dAC. However, the DMT-dAC peak did decrease, which was potentially due to triazine ring opening, the product of which could not be detected via HPLC.

After further column purification, the DMT-dAC was successfully isolated (Figure 5.2). Subsequent stability analysis on the twice purified analogue showed that after 24 hours, a noticeable peak can be detected which could correspond to a very small amount of dAC that may have dissociated during the incubation (Figure 5.3).

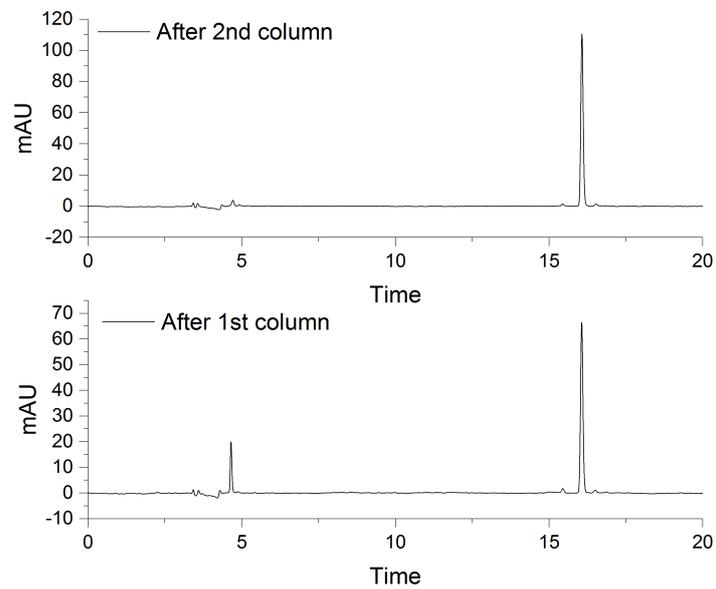


Fig. 5.2 HPLC traces reflecting the purification steps of DMT-dAC

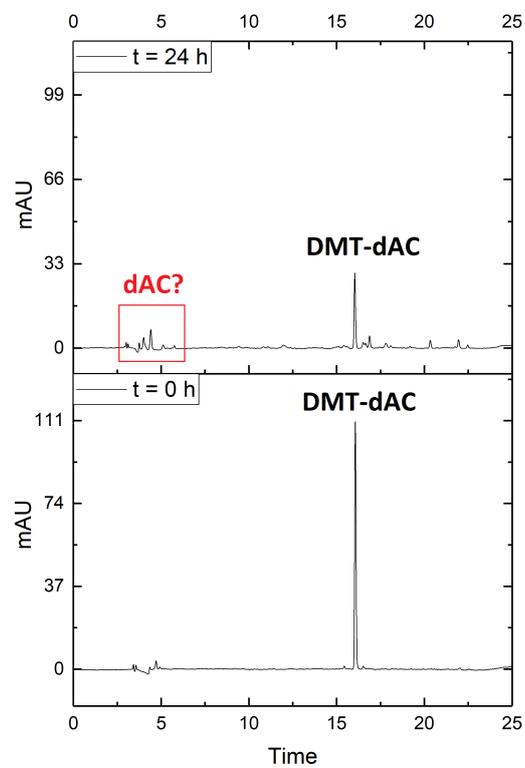


Fig. 5.3 HPLC Stability analysis of twice purified DMT-dAC, 97

Given that it could not be confirmed whether the DMT-dAC analogue had superior hydrolytic stability compared to the photocaged analogues, it was decided not to test this analogue further in cell biological experiments. The hydrolysis of the DMT-conjugated analogue could be explained due to the presence of methoxy groups on the phenyl groups. The presence of the methoxy group may have increased the electron density of the functional group, making the DMT group a better leaving group and more prone to dissociation.

5.4 Suggested future work

Future work in this section could involve exploring alternative functional groups that can be linked to 5'-OH of dAC. Potential linkers could include an ether or other hydrolytically stable bonds. Once a non-hydrolysable analogue of dAC is synthesised and characterised, stability analyses need to be conducted using LC/MS (as described in Chapter 4), rather than HPLC. It is essential to identify an analogue that is not prone to hydrolysis in aqueous media at 37 °C.

After a hydrolytically stable analogue is identified and established, cell experiments can be conducted to determine whether the non-hydrolysable version of dAC is biologically active or not (ie. degrade DNMT1). To determine that, western blot analyses can be performed to measure the level of DNMT1 in response to the analogue.

If the non-hydrolysable analogue is able to reduce DNMT1 level, that would suggest that any 3'- or 5'-modified version of dAC could induce demethylation even in the absence of replication, since these modifications block the incorporation of dAC into the DNA. This would also suggest that phototags on these positions would not have a activity masking effect. However, if these non-hydrolysable analogues do not reduce DNMT1 level, that suggests that 3'and 5'modifications (including photocaging) are able to block dAC bioactivity.

Chapter 6

Conclusion

6.1 Summary and aim of the work

DNA methylation plays an essential part in maintaining and regulating key cellular events. Aberrant changes in DNA methylation patterns such as hypermethylation or increased methylation can lead to the development of cancer. DNA methylation inhibitors such as azanucleosides (AC, dAC) can reverse hypermethylation by covalently blocking DNMT1 enzyme. However the therapeutic use of these analogues can lead to severe side effects (e.g. cytotoxicity) due to their non-specific nature. It would be desirable to be able to target the activity of these agents to the cancerous site of interest, in order to reduce potential side-effects. There is a significant need for epigenetic tools that can be switched on and off “on demand”, both for therapeutic and research applications. For research applications, these agents can serve as tools to deepen our understanding of epigenetic modifications, which would be invaluable to the scientific community.

The aim of this work was to develop optically controlled small molecule methylation tools to target epigenetic modifications, more specifically DNA methylation. A series of novel photocaged azanucleoside analogues were synthesised and characterised for the purpose of light-dependent demethylation. It was established that caging position strongly affected the photochemical and stability characteristics of the caged compounds. Photocaged azanucleosides modified at the exocyclic $-NH_2$ experienced significant dark instability. This was due to the electronic properties of the triazine nucleobase that led to the dissociation of the phototag. Placing the phototag on the 5'-OH greatly

mitigated this effect, although the light-activation rate of this analogue suffered from slower uncaging rates. The 3'-modified analogue (3'-DEACMOC-dAC) was the most promising candidate due to its rapid deprotection rates and significantly improved stability compared to the N-caged analogue.

Photocaged azanucleosides are biologically inert forms of azanucleosides that can be activated upon light irradiation. Cell biological studies on SaOS2, T24 and HEK293T cells showed that 3'-DEACMOC-dAC induced light-dependent demethylation, but the effect was limited to lower concentration treatments ($< 0.5 \mu\text{M}$). At higher concentrations, significant demethylating activity was observed even in the absence of light, suggesting that active dAC may have been released. These findings were corroborated by Western blot analyses and cell viability assays.

Further investigation into the stability profile of 3'-DEACMOC-dAC using LC-MS/MS analysis suggested that the mechanisms contributing to the degradation of this analogue were the triazine ring opening ($t_{1/2} = 11.5 \text{ h}$) and carbonate link hydrolysis ($t_{1/2} = 5 \text{ h } 50 \text{ min}$) when the compound is incubated under physiological conditions. The latter event also resulted in the release of 8% active dAC from photocaged dAC over a 6 hour incubation period, which likely caused the effects observed in the cell biological experiments.

6.2 Impact and significance of the work

6.2.1 Significance of the work

This work describes the design and synthesis of chemical tools, more specifically photocaged azanucleoside analogues that can achieve light-dependent demethylation within the cellular environment. This is the first work in the field to describe the photocaging strategy of azanucleoside analogues, which will be a valuable addition to the limited amount of literature reports on the chemical modification of azanucleosides.

6.2.2 Related work described in the field

Modified azanucleoside analogues

There were a number of efforts in the past 20 years to design safer and less cytotoxic analogues and prodrugs of azanucleosides that can still achieve significant demethylation, but with reduced or no side effects.⁸⁷ The most successful clinical candidate to date is SGI-110, a dinucleotide prodrug of dAC, in which the active dAC is linked to deoxyguanosine (dG) via a phosphodiester bond. The analogue's differentiated pharmacokinetic and pharmacodynamic profiles offer improved clinical efficacy over existing hypomethylating agents, not only in patients with haematological malignancies, but the drug also demonstrated efficacy in the treatment of solid tumours.²⁵⁶ The agent has been tested in bladder and colon cancer cells and is currently in phase III trials. While there is clear evidence that the analogue offers improved safety profile, its activity and efficacy relies on the presence of a phosphodiesterase enzyme that is required for the release of active dAC.⁸⁷

Another, similarly promising analogue was the N-modified 4-NPEOC-dAC, which was originally developed by Byun et al to circumvent the metabolic instability of dAC due to the inactivation by cytidine deaminase (CDA) enzyme.¹⁶⁶ This analogue was synthetically modified at the exocyclic amine by addition of a chloroformate version of the functional group (4-NPEOC), which then rendered the dAC physiologically inactive. The synthetic protocols applied in this work served as a basis and initial starting point for this thesis. The demethylation activity of the prodrug was proven by cell biology experiments.¹⁶⁶ These experiments showed that NPEOC-dAC significantly decreased global DNA methylation by reversing hypermethylation and reactivate the expression of TSG and ID4 genes. However, the DNA methylation ability of the analogue was found to be specific to liver cancer cell lines.

Similarly to SGI-110, the activation of this analogue relies on an enzyme, in this case Carboxylesterase 1 enzyme (CES1), which is unevenly distributed in different tissues and is more abundant in liver cells, but may be found in lower amounts in other tissues such as ovarian cells. In addition, the analogue also had a 3 day delay in effect and less toxicity was observed compared to dAC. The low potency and delayed effect was due to the inefficient and slow conversion of NPEOC dAC to the active dAC. The fact that NPEOC-dAC is dependent on the activity of CES1 enzyme for its metabolisation to dAC has limited the clinical development of this prodrug, because the expression of

CES1 is variable in different tissues and also the enzyme may not be 100% efficient in converting NPEOC-dAC to dAC.

In summary, two promising prodrugs of dAC have been developed so far (one in clinical development and one in pre-clinical stage). Both compounds rely on enzyme activation to release active dAC. While it was demonstrated that these analogues have superior safety and toxicity profiles compared to dAC, the activation is somewhat a limiting factor as the presence of the enzymes in tissues cannot be controlled by external factors.

On the other hand, activation of photocaged azanucleosides is controlled by light, which can be easily manipulated externally, for example by changing the location, timing and intensity of the light source. Therefore photocaged azanucleosides are a valuable addition to the current selection of azanucleoside prodrugs, as they can be activated under precise spatiotemporal control, which could potentially allow the release the exact amount of active compound in a given space at a given time.

Photocaged nucleoside analogues to study epigenetic processes

Photocaging of nucleoside analogues is a well-studied area and there are many instances, where photocaging strategy was developed to understand the details of biological processes, mechanisms that involve nucleoside analogues, which are comprehensively discussed in a review by Tang et al. For example caged-ATP (for studying the energy-dependent steps of metabolic reactions) and caged cyclic nucleotide monophosphates (studying signal transduction pathways where secondary messengers play a key role).

To date, very few photocaged analogues have been synthesised to study epigenetic processes, but this is likely to change in the near future due to the increasing interest in understanding the mechanism of epigenetic modifications and its impact on the development of cancer. The synthesis of one of the first analogues - photocaged 5 hydroxymethylcytosine was recently reported by a research group in the Czech Republic.²⁵⁷

In their study, two types of photocaged derivatives of 2'-deoxy-5-hydroxymethylcytosine dNTPs were synthesised. One was protected by 2-nitrobenzyl (one of the most simple photocages) and the other one had a 6-nitropiperonyl group to mask its activity. Both analogues were tested for enzymatic incorporation into DNA and photochemical deprotections. While the nitropiperonyl group was efficiently deprotected by UV irradiation, the nitropiperonyl group was more suitable with visible light at 400-425

nm. The authors argued that the NB-caged analogue still has a good potential for photocaging of DNA containing 5 hmC and some biological applications including studies in order to better understand the role of 5 hmC in the regulation of transcription, genomic stability and active demethylation.²⁵⁷

The photocaging of another base, 5 hmU was reported by the same group.²⁵⁸ 5 hmU is a less known epigenetic mark as it is usually found in certain bacterial species. Similarly to the 5 hmC analogue, photocaged versions of this analogue were synthesised the probe their ability to be incorporated into DNA and their photochemical deprotection was checked. Apart from these two analogues, there are no other photocaged small molecules that can selectively target DNA methylation processes.

Light sensitive biological tools to study epigenetic processes

On the other hand, biological tools that can finely manipulate epigenetic modifications even with the control of light are rapidly emerging in the last few years with the recent development of the CRISPR-CAS system. An example of such system was reported by Polstein and coworkers in 2015.¹⁷⁷

In their work, the authors engineered a light-activated CRISPR-Cas9 effector (“LACE”) system that could induce transcription of endogenous genes in the presence of blue light. They fused the light-inducible heterodimerizing proteins (CRY and CB1) to a transactivation domain and the catalytically inactive dCas9 respectively. The resulting LACE system can easily be directed to specific DNA sequences for the dynamic regulation of endogenous genes.¹⁷⁷ The light-inducible CRISPR-CAS system enabled spatial and temporal control of endogenous gene activation and cell behaviour (e.g. increase in mRNA levels upon illumination) as claimed by the authors.

While the CRISPR-Cas system offers a versatile platform that can be tailored to any DNA sequence, the gene activation for the LACE system is reversible through removal of illumination. In the case of the photocaged-dAC, the activation is non-reversible. Azanucleosides are not targeting specific DNA sequences, but rather they are “global demethylators”. Therefore photocaged-azanucleosides (although could be switched on and off with the control of light) would therefore offer a less targeted approach compared to the light inducible CRISPR-CAS system in terms of locus-specific modifications, but its effect is non-reversible (at the time of the illumination).

6.3 Limitations of the work

The current work has three main limitations. The first limitation was the chemical instability of the compound. The second limitation, which is partly related to the compound's instability, is the setup of the cell biological experiments. And finally, the stability assays conducted with mass spectrometry also had a number of drawbacks that could potentially be addressed in future experiments. In this section, each of these points will be discussed in detail and how these limitations could be improved in future studies.

6.3.1 Limitation 1: Hydrolytic Instability

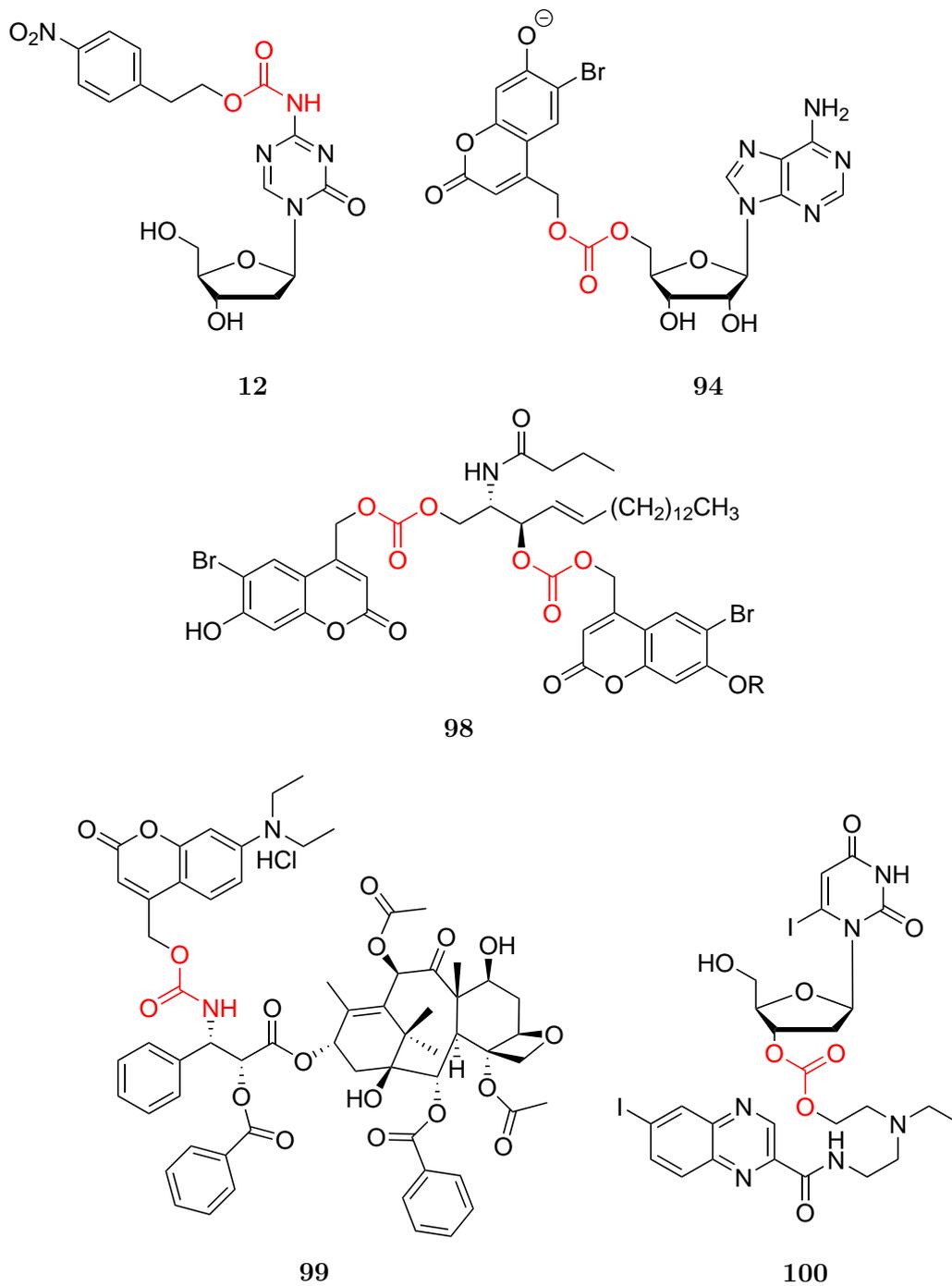
Unfortunately the hydrolytic instability of the carbonate bond linking the photocage and the dAC limits its cell biological applications of photocaged azanucleosides as it resulted in significant activity of the compound even in the dark.

Literature survey of related compounds and their reported instabilities

In light of this limitation, literature reports of analogues with similar structures were investigated in order to determine if any stability issues were reported and how the authors addressed these limitations. Similar structures included nucleoside analogues containing a carbamate or carbonate linkers (they could be photocaged or not photocaged). Alternatively, any other photocaged analogues, where the photocage is connected via a carbamate or carbonate linker to the active compound were considered. Five analogues were identified from the literature search, the structures of which are shown in Scheme 6.1.

Of the five, three are nucleoside analogues (**12**, **94**, **100**), one is an azanucleoside analogue (**12**) and three are photocaged compounds (**94**, **101**, **102**).

The first analogue (**12**), which is the most closely related to photocaged azanucleosides is 4-NPEOC-dAC, which was already introduced in the previous sections. This analogue is not a photocaged azanucleoside as the 4-NPEOC group cannot be removed by light. But rather, 4-NPEOC-dAC is a prodrug of dAC that relies on enzymatic activation.¹⁶⁶ The NPEOC group is linked to dAC via a carbamate bond. There was no mention in the study regarding the stability or instability of the compound and the cell experiments suggested that there was no dark instability observed. This is intriguing



Scheme 6.1 Carbonate/carbamate containing nucleoside and non-nucleoside analogues

as in our experience, N-modified dAC analogues had poor hydrolytic stability, which was confirmed by HPLC analysis at physiological conditions.

The second analogue (**94**) is a photocaged nucleoside analogue, coumarinyl adenosine (5'-Bhmoc-adenosine).¹⁹¹ In this analogue, the coumarin photocage is linked to the nucleoside via a carbonate linker at the 5'-position of the nucleoside. Apart from the adenosine, other compounds such as galactose were also caged to coumarin photocages via a carbonate linker in the study.¹⁹¹ According to the authors, these compounds show "modest to good hydrolytic stabilities". Their half-lives in the dark was reported to be 467 h for the photocaged nucleoside analogue. However, the temperature conditions for the stability analysis were not reported, likely that they were conducted at room temperature rather than 37 0 °C.

The third analogue (**98**) is a BHC-ceramide, a photocaged small molecule and not a nucleoside analogue. In this analogue, the phototag is linked to the ceramide active molecule via a carbonate linker. Similarly to the previous compound, the half-life was reported to be very good (30 hours).²¹⁷ However, it was stated that the temperature at which stability was measured was room temperature, which cannot be considered physiologically applicable.

The fourth analogue is "phototaxel" (**99**), a water-soluble photoresponsive prodrug of the anticancer agent paclitaxel. In the study, the authors made a carbamate and carbonate type versions of phototaxel and measured their stabilities at physiological conditions.²⁵⁹ The carbamate version of phototaxel was stable in the dark prior to activation, whereas the carbonate version showed poor aqueous stability in the dark. More specifically, only 20% of paclitaxel was released in PBS (pH 7.4, 37 0 °C) after 8 hours of incubation and only 18% of the prodrug remained when the analogue was stored at 4 0 °C for 13 days. In contrast, the carbamate analogue was stable under physiological conditions as well as under alternative storage conditions. Therefore, for latter studies the authors chose the carbamate analogue for further detailed study.²⁵⁹ This study is an example of where the instability of the carbonate-type analogue was identified and addressed by the authors, by switching the carbonate to a more stable carbamate linker.

In the final analogue (**100**), the antitumour agent 5-iodo-2'-deoxyuridine nucleoside was conjugated to a melanoma carrier molecule via varying linker structures for potential intratumoural specific drug release.²²⁰ The authors explored various linker types including ester, acetal, carbonate and carbamate. The stabilities of each analogues (with different linker types) were tested in various physiologically relevant conditions

such as PBS (pH 7.4), cell culture medium (MEM) and human serum, as well as acidic buffer (pH 5.0). As demonstrated by the Figure 6.1, the analogue with the carbonate linker had a poor aqueous stability. After 6 hours of incubation in PBS, MEM and human serum, only 60%, 50% and 40% of the carbonate analogue remained in the solution respectively. In acidic conditions, the stability was just as poor as in human serum. The authors reported an even poorer stability profile for the ester analogue, but the carbamate conjugate on the other hand was reported to be stable over 3 days in PBS and MEM (data was not shown in the article).²²⁰

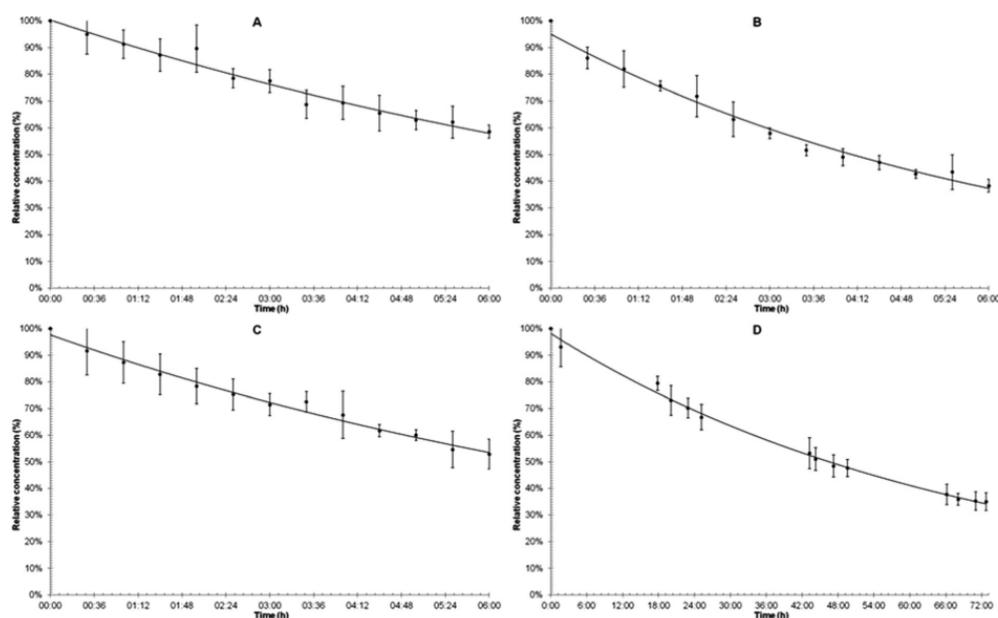


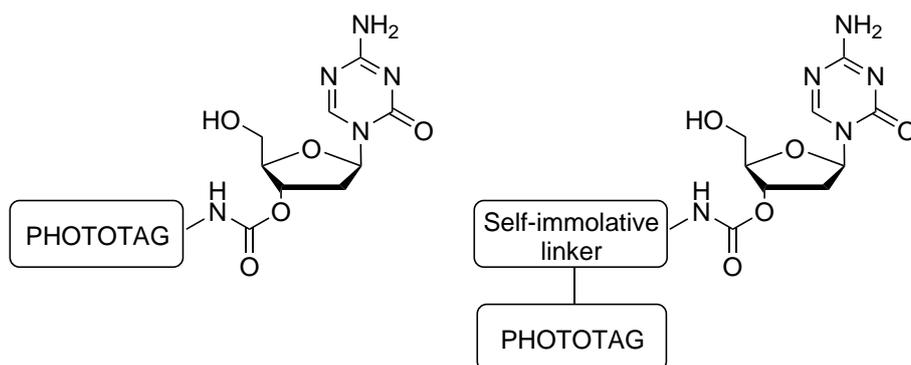
Fig. 6.1 The reported stability profiles by Andre and coworkers on carbonate linked 5-iodo-2'-deoxyuridine **100** to a melanoma carrier molecule in (A) PBS buffer (B) human serum (C) cell culture media and (D) acidic buffer. The profiles confirm that carbonate linked analogues have poor aqueous stability under physiologically relevant conditions.²²⁰

These results are in alignment with our findings. In our stability experiments, the stability of the carbonate conjugated cytidine analogue was comparable to that of analogue **100**. After 6 hours of incubation in HEPES at pH 7.2, around 53% of the photocaged compound remained in solution, which is slightly lower to what is reported by the authors of this study (60%). Based on these findings, it is likely that the aqueous stability of the photocaged azanucleosides would be even poorer in cell culture media and human serum. Instead of modifying the structure of the carbonate linked analogue, the authors confirmed that the carbamate linked analogue has a superior stability profile and therefore should be used in further cell biological experiments.

Based on the five literature studies reporting on the stability of carbonate and carbamate linked analogues, it can be concluded that none of the carbonate linked analogues had good aqueous stability under physiological conditions. In the two examples where the authors claimed good stability and long half-life, the stability studies were conducted under non-physiological conditions. In the studies, where the aqueous instability was evaluated under physiologically relevant conditions, the half-lives obtained were very similar to the ones obtained from our study. There was only one way the authors addressed this instability issue and it was simply using the carbamate analogue of the same compound instead of the carbonate analogue for further biological studies.

Alternative structures with improved hydrolytic stabilities

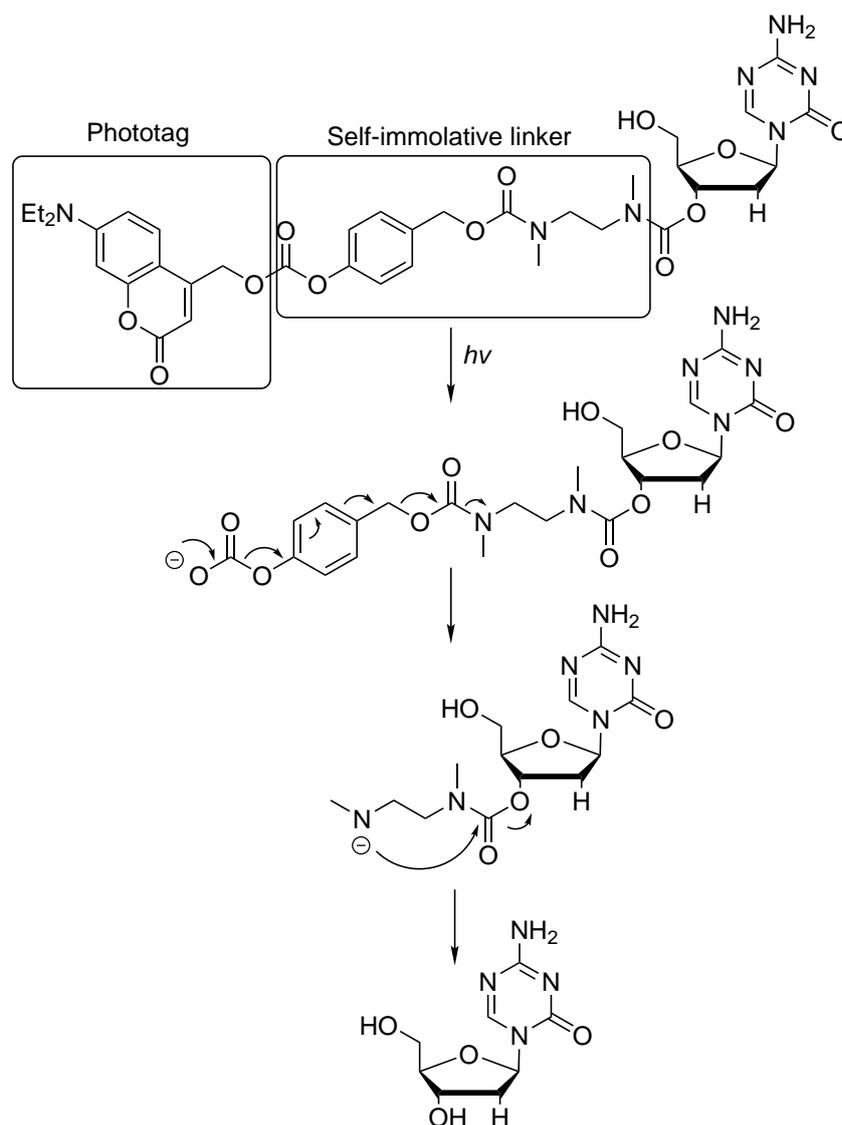
In the case of photocaged azanucleoside, replacement of the carbonate linkers with carbamate bonds would require the modification of the photocage. The position of the photocage is fixed to the 3'-position of dAC, due to the limitations in stability and deprotection kinetics as observed for the N'- and 5'-caged dAC analogues previously. This requires the Oxygen atom to be part of dAC (3'-position) and the Nitrogen atom to be part of the photocage. Therefore, either the part of the photocage that links to dAC must contain a Nitrogen atom or there should be an additional linker between the photocage and the carbamate linker (e.g. a self-immolative linker) that enables this connection (Scheme 6.2).



Scheme 6.2 Alternative caging strategy via carbamate linkers

Self-immolative linkers

Self immolative linkers are special types of linkers. They can undergo a controlled cleavage mechanism upon a certain kind of trigger such as photouncaging, as well as enzymatic, nucleophilic or redox processes.²⁶⁰



Scheme 6.3 Self immolative linker strategy

In the case of photouncaging, this linker forms a scissile bond with the protecting group and a stable bond to the target molecule (in this case dAdC). The latter bond becomes labile upon removal of the protecting group, resulting in a rapid disassembly of the three components.

Self-immolative linkers are becoming increasingly common in conjugate systems for drug delivery.²⁶¹ An example of a photocaged-dAC structure containing a self-immolating linker and its proposed decomposition upon UV irradiation is shown in Scheme 6.3.

Self-immolative breakdown occurs through a cascade of electronic elimination processes and it is driven by an increase in entropy and the irreversible formation of thermodynamically stable products (in this case CO₂, 5-membered ring and the photocage).

It is hypothesised that the insertion of a self-immolative linker between the DEACM photocage and dAC would ensure rapid removal of the photocage and the linker upon light irradiation. Furthermore, the self-immolative linker also allows DEACM phototag to be attached to dAC via a carbamate linker, which is expected to significantly increase its hydrolytic stability as suggested by literature.

6.3.2 Limitation 2: The setup of cell biological experiments

The aim of the cell biological experiments was to demonstrate how the DNA methylation levels of the cells were altered upon various treatments (dAC, photocaged-dAC and irradiated photocaged-dAC). One of the most challenging steps of the setup of the cell biological assays was to identify a robust methylation assay that can determine the change in DNA methylation levels of the treated cells at a relatively rapid rate. Given the lack of sensitivity of the ELISA assays, LC-MS was identified to be the most suitable technique to reflect even small changes in DNA methylation levels upon treatment.

Unfortunately it was not possible to set up the LC-MS assay "in house", given the lack of time for optimisation and access to specialised equipment. An alternative laboratory in Cambridge was selected to conduct the LC-MS assays, where an already established and validated protocol was used to measure DNA methylation levels using LC-MS. For each measurement cycle, a limited number of samples could be submitted and due to the limited access to the equipment, it often took between 2 weeks to 1 month to receive the methylation results for the samples that have been submitted. Given these limitations, it was not possible to send a large number of samples at once.

The cell line screening described in Section 3.4.1 was one of the examples where this resource limitation had a potential impact. In the screening, four different cell lines were tested for their sensitivity to dAC treatment, at 5 different concentrations. Ideally, the treatments should have been carried out in triplicates, which would have resulted

in 60 samples. But at the time, only one replicate was produced for each treatment condition and cell line, which resulted in 20 samples. It is recognised that this is a potential limitation of the work and having triplicates of the samples would have provided higher level of confidence in the results obtained. However, because the following readings with the selected cell lines (SaOS2, T24 and HEK293T) always included triplicate measurements, it was not considered essential at the time to repeat the measurements for the cell line screening.

Similarly for the treatment optimisation, potentially more concentrations could have been explored (for example for SaOS2 cells, between 0 and 0.1 μM to determine a concentration where SaOS2 cells still respond to dAC in a concentration dependent manner). But given the limitations described above, these measurements were not carried out.

Another limitation of the cell experimental setup was the treatment duration. Initially, a four day treatment regimen was trialled, where the cells were treated with dAC every 24 hours. It was then decided to switch to a single treatment which only lasted 24 hours. The switch was performed, because a notable level of toxicity was observed at the highest concentration dAC treatment (1.5 μM), which affected the amount of DNA that could be extracted from the cells.

With the new treatment regimen, significantly greater amount of DNA could be obtained, however it is also recognized in hindsight that 24 hours is not a long enough time for dAC to be fully incorporated into the DNA. Therefore a lower concentration (e.g. below 0.5 μM) treatment for longer duration (for 3 days and potentially even longer) potentially would have better reflected the effect of dAC on DNA methylation. In literature, dAC treatments are usually conducted for 72 hours in different cell lines.²⁶²

6.3.3 Limitation 3: Stability Experiments

Another set of limitations of this thesis were concerning the stability assays. The aim of the stability assays was to gain additional (qualitative and preferably quantitative) insights on the potential mechanisms that contributed towards the hydrolytic instability of the photocaged-dAC compound that resulted in unwanted biological activity of the compound in the dark.

The stability analysis using mass spectrometry enabled the detection of dAC, caged-dAC and photocage in various conditions (water, HEPES buffer at room temperature and 37 °C) over a period of 6 hours. The current setup enabled the collection of limited amount of data points (in time) and it allowed the detection of limited number of compounds (only dAC, caged-dAC and photocage, but none of the potential intermediates inbetween).

Limited amount of data points

The sample collection interval of the stability analysis was limited to a minimum of 30 minutes as this amount time was required for the full separation and individual detection of the caged-dAC compound and the DEACM photocage on the liquid chromatography column.

For the stability measurements at 37 °C, an endpoint at 6 h was chosen due to the limited availability of LC/MS machine and requirement to manually transfer samples to the LC/MS vials. The reason for manual transfer was because the samples were incubated in a separate container (a heating block) from the LC/MS sample holder. Prior to each injection, 200 μL aliquot of each sample was placed in an LC/MS vial and then sample holder. After the injection (of approximately 10 μL), the remaining sample was returned to the heating block to maximise the available liquid volume for subsequent stability analyses.

Detection of limited number of compounds

The current setup only enabled the detection and quantification of three compounds: dAC, caged-dAC and the photocage. Only these compounds were available in a pure and isolated form for initial calibration, which was preceded by MRM (multiple reaction monitoring) optimisation using the triple quadrupole mass spectrometry. Subsequently, LC/MS only allowed the quantification of compounds that underwent MRM optimisation and calibration.

The ability to monitor the levels caged-dAC gave an exact indication of the extent of hydrolytic instability of the compound, demonstrated by its rate of disappearance. The concomitant release of dAC (8 % of starting caged-dAC concentration) measured also reflected the amount of biological activity that was observed in the cell experiments (ie. equivalent to dAC activity at approximately tenth of the concentration).

Need for additional reference compounds for stability analyses

It was possible to confirm through the series of stability experiments (with photocaged-dAC, dAC, but also photocaged-dC) that there were two main sources of the hydrolytic instability: triazine ring opening and carbonate bond dissociation. However, it was not possible to quantify the presence of any degradation intermediates and byproducts (such as the triazine ring opened forms or the hydrolysed/dissociated carbonate dAC intermediate) due to the lack of pure and isolated form of these compounds for MRM optimisation and calibration.

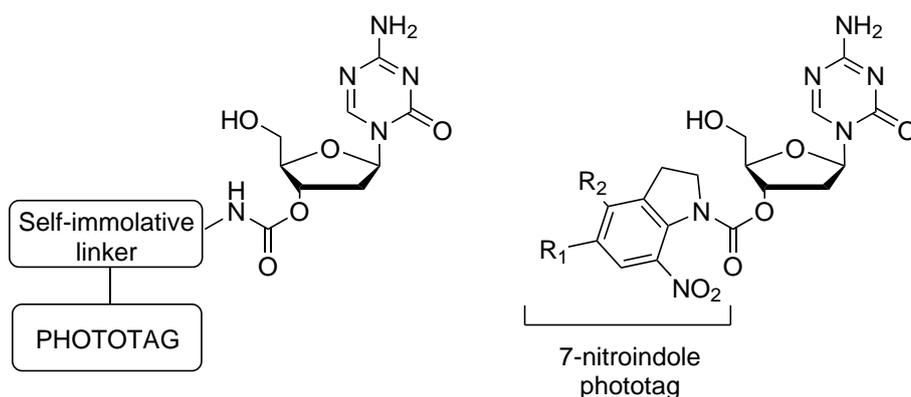
It was particularly challenging to measure the rate of decarboxylation as there was no indication of the level of carbonated dAC. The only available measure was the appearance of the photocage (step prior the appearance of carbonated dAC) and the appearance of pure dAC (step after the carbonated dAC). Similarly, the rate of ring opening was also difficult to measure in the absence of a standard solution of the ring opened form. Independent measurements of dAC triazine ring opening and photocaged-dC stability did help stability measurements, however we had to assume that the behaviour of these bonds are similar within photocaged-dAC.

A more comprehensive and quantitative stability study would require additional reference compounds (for calibration), such as the ring opened triazine form or the carbonated dAC intermediate. The ring opened form is possible to isolate via preparative HPLC from a 24 h old aqueous solution as described by Beisler et al.¹⁵⁴ However, the carbonated dAC would not be possible to isolate as the compound itself is labile.

In summary, a number of assumptions were stated and tested during the stability studies, but only qualitative and limited quantitative conclusions could be drawn due to the small number of compounds that could be monitored (ie. no degradation intermediates and by-products). Further improvements can be done by including more reference compounds in order to gain further insights on the individual mechanistic steps that lead to the triazine ring opening or the carbonate bond dissociation.

6.4 Suggested future work

Based on our observations and the discussion of the current work's limitations, we suggest that the stability of the caged analogues could be improved by replacing the carbonate bond with a carbamate linker. Given the instability of the exocyclic position in azanucleosides and poorer uncaging rates observed at the 5'-position, the most ideal position for the carbamate-linked phototag would be at the 3'-position (Scheme 6.4).



Scheme 6.4 Suggested analogues with improved hydrolytic stabilities

This would require a phototag with an amine handle. Unfortunately, the DEACM phototag does not have an amine handle. It is however possible to keep the DEACM phototag by the introduction of a self-immolative linker as described in the previous section. Alternatively, a different phototag can be used. The 7-nitroindole phototag is one of the rare phototags that has an amine handle, but its photochemical properties are considerably poorer than coumarin phototags' used in this work.^{181,263}

Further work could either involve the improvement of the photochemical properties of the 7-nitroindole phototag or the photochemical investigation of the combination of DEACM phototag and self-immolative linker. All of the future photochemical investigations should be conducted in physiological buffer instead of water.

Stability experiments should ideally be conducted prior to any cell biological experiments to ensure that the compound has good aqueous stability even in the dark. Similarly to the DEACM-dAC analogue, a dC analogue (that does not have triazine ring instability) should also be synthesised so that the carbamate bond stability can be probed individually.

In terms of the cell biological experiments, treatment duration should be extended to 72 hours from 24 hours and the treatment concentrations should be focused on the

lower range ($< 0.5 \mu\text{M}$). LC/MS measurements should be kept for DNA methylation analysis, but it should be ensured that measurements are conducted in triplicates.

Azanucleosides provided an important first step towards the development of light-tunable small molecule methylation tools. In order to achieve greater therapeutic impact, DNMT inhibitors other than azanucleosides should also be explored in the future. Currently, azanucleosides are the most potent global hypomethylators. With the rapid development in the field of epigenetics, the discovery of new small molecule DNMT inhibitors with greater potencies will provide additional platforms for photocaging and controlling methylation levels with the precision of light.

Appendix A

Experimental

A.1 General experimental

A.1.1 Chemicals and solvents

All chemicals and solvents were purchased from Sigma Aldrich, Fischer Scientific, Alfa Aesar, Acros Organics and used as received unless otherwise stated. 5-azacytidine and 5-aza-2'-deoxycytidine were purchased from Carbosynth Ltd.

A.1.2 Chromatography

All reactions were monitored by thin layer chromatography using Merck 60 F254 silica gel coated plates. Visualisation was achieved by UV irradiation at 254 nm or 365 nm. Flash column chromatography was carried with silica gel (43 – 60 μm) supplied by Merck.

A.1.3 Spectroscopy

Nuclear Magnetic Resonance (NMR)

^1H NMR spectra were recorded at 500 MHz and 600 MHz and ^{13}C NMR at 75 MHz, 100 MHz, 125 MHz and 150 MHz on Bruker AMX500 and AMX600 at 25 °C as described below. All chemical shifts (δ) for ^1H and ^{13}C are quoted relative to residual signals

of the solvent on the parts per million (ppm) scale. The multiplicity of the signal is designated by the following abbreviations: m=multiplet, s=singlet, d=doublet, t=triplet, q=quartet. The abbreviation br refers to a broad signal. Coupling constants (J values) are reported in Hz and refer to $J_{\text{H-H}}$ couplings unless otherwise stated.

Infrared Spectroscopy (IR) and melting point analysis

Infrared spectra were obtained on a Perkin Elmer Spectrum 100 FTIR Spectrometer operating in ATR mode. Melting points were measured with a Gallenkamp apparatus and are uncorrected.

UV-Vis

UV analysis was performed using Varian Cary 300 Bio UV-Vis Spectrophotometer and a quartz cuvette with a path length (l) of 1 cm. All solutions were prepared at a concentration (c) of 50 μM in water unless otherwise stated. Molar absorption coefficients values (ϵ) were calculated according to the Beer Lambert Law.

$$\text{Absorption} = \log_{10} \frac{I_0}{I} = \epsilon lc \quad (1)$$

LCMS, HRMS

ESI-MS analysis was performed using a Waters Acquity Ultra performance LC-MS system equipped with an Acquity UPLC BEH C_{18} column (50 x 2.1 mm, 1.7 mm beads). CI-MS analysis was performed on a Thermo Finnigan MAT900 XP equipment by Dr Kersti Karu, UCL Chemistry.

A.1.4 HPLC

HPLC analysis was performed using reverse phase Varian ProStar system with a Model 210 solvent delivery mode and a Model 215 dual wavelength detector. The UV absorbance was monitored at $\lambda = 254$ nm. Semi-preparative HPLC was carried out using a Supelco C₁₈ column (250 x 21 mm, 10 μ M) column with an elution system of 5-95% ACN in water over 45 min at 5 mL.min⁻¹ flow rate. Analytical HPLC analysis was performed using an Agilent Eclipse C18 column (250 x 4.6 mm, 5 μ m), with an elution system of 5-95% ACN in water over 20 min at 1 mL.min⁻¹ flow rate.

A.1.5 Photouncaging experiments

Photouncaging was carried out inside a UV Viewing Cabinet (UVP, #95-0072-01) equipped with a handheld UV lamp (Mineralight, #UVGL-58) at a measured power density of 145 μ M/cm². The photouncaging solutions were prepared at 100 μ M in water from frozen 10 mM DMSO stocks of photocaged azanucleosides. 500 μ L of the freshly prepared photouncaging solutions were placed in a 24 well tissue culture plate (BD Falcon, #353047) and positioned 20 cm from the UV lamp inside the UV Viewing cabinet. Irradiations were carried out at $\lambda = 365$ nm, unless otherwise stated. 20 μ L aliquots were taken every 30 min and the concentration of the caged compounds and the liberated azanucleosides were measured using HPLC.

A.1.6 Quantum yield measurements

Uncaging reactions are typically carried out using xenon or mercury lamps. Quantum yield measurements for photouncaging reactions are conventionally determined using a combination of potassium ferrioxalate actinometer (to measure the amount of photons absorbed) and HPLC (to measure the ratio of molecules converted from caged to uncaged state). The photochemical quantum yield is defined as the ratio of caged molecules converted to the amount of photons absorbed that was measured by the actinometer.

For our uncaging setup, two adjustments were made to the conventional methods described in literature. The first adjustment was the light source used for the irradiation experiments. Instead of conventional xenon and mercury lamps, a benchtop UV lamp was used to irradiate the compounds. The benchtop UV lamp was found to be a

convenient and compatible choice for the chemical photouncaging experiments as well as the cell biological assays that required photoirradiation of cell cultures. The second adjustment was the measurement tool for the photon influx. In the absence of an actinometer equipment at the time, an alternative method was employed that was already used to measure quantum efficiency of photocatalytic reactions.^{264,265} In this method, the amount of photons absorbed was measured using a spectrophotoradiometer (UVX Radiometer) that was positioned in the place of the uncaging solution. The ratio of the molecules converted was determined via HPLC, following literature reported protocols.

There are several limitations of this method and it is important to point out that the “yield” calculated here is the *formal quantum efficiency* rather than the quantum yield. The spectrophotometer measures that number of incident photons - that is the number of photons that reach the solution. The limitation is that the actual number of photons absorbed by the solution is not known using this measurement, and it is tricky to determine as some photons may get lost via radical chain reactions participating in the degradation process. The formal quantum efficiency / apparent quantum efficiency gives an idea of the efficiency of the uncaging process, however it relies on the measurement of incident photons rather than the photons absorbed. Therefore, the values calculated are likely an underestimation of the actual quantum yield values.

A.1.7 Quantum yield calculations

The quantum yields, Φ , defined as the ratio of caged molecules converted to the amount of photons absorbed were estimated based on the equations below.

$$\text{Photon Flux (photons.cm}^{-2}\text{s}^{-1}) = \frac{\text{Irradiance (mW/cm}^2\text{)}}{\text{Energy (J)}} \quad (2)$$

$$\text{Quantum Yield } \Phi \text{ (molecules.photon}^{-1}\text{)} = \frac{\text{Molecules destroyed (molecules.cm}^{-2}\text{s}^{-1}\text{)}}{\text{Photon Flux (photons.cm}^{-2}\text{s}^{-1}\text{)}} \quad (3)$$

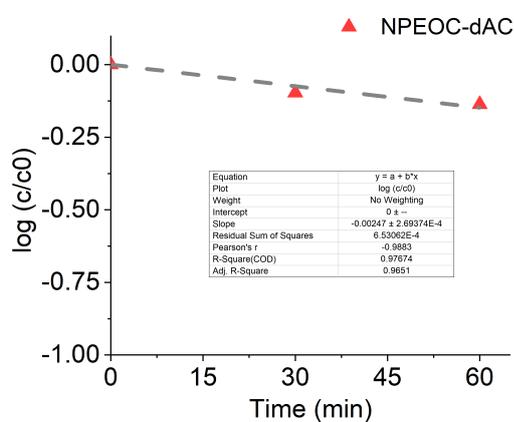
The calculated photon flux at $\lambda = 365$ nm at an irradiance of $145 \mu\text{W}/\text{cm}^2$ was 2.64×10^{14} photons.cm⁻²s⁻¹. Table A.1 summarises the calculated quantum yield values of the photocaged analogues.

Photocaged analogue	Rate of decay (s ⁻¹)	Molecules destroyed (molecules cm ⁻² s ⁻¹)	Formal Quantum Efficiency $\Phi_{365\text{nm}}$
mono-NPEOC-dAC, 60	8.33E-05	1.32E+12	4.93E-03
mono-DMNPEOC-dAC, 64	6.67E-05	1.05E+12	3.94E-03
N-DEACMOC-dAC, 78	1.03E-03	7.11E+12	6.11E-02
5'-DEACMOC-dAC, 84	4.83E-04	3.34E+12	2.88E-02
3'-DEACMOC-dAC, 93	1.50E-03	1.03E+13	8.84E-02

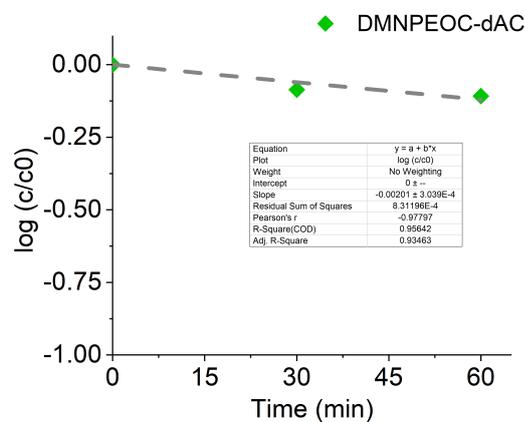
Table A.1 Quantum yield calculations of photocaged analogues

A.1.8 Rate calculations

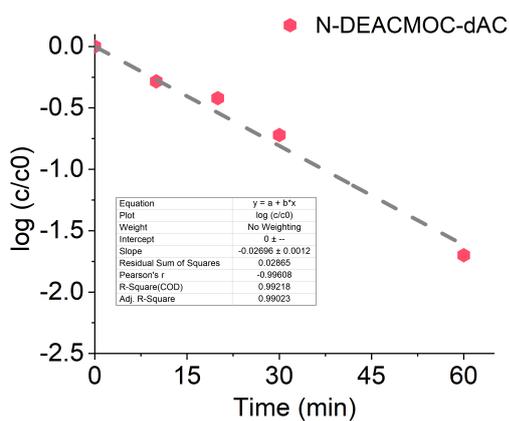
Photouncaging rates were estimated from linear regression of the 30-40% steps of the area curves extracted from zero order kinetics (as shown in the Figure A.1). Log of measured concentration and initial concentration ratio was plotted against time and the measured slope of the linear function was taken as the corresponding rate of the decay.



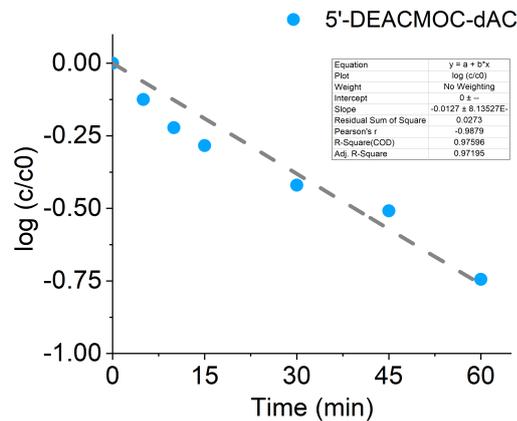
(A) mono-NPEOC-dAC, 60



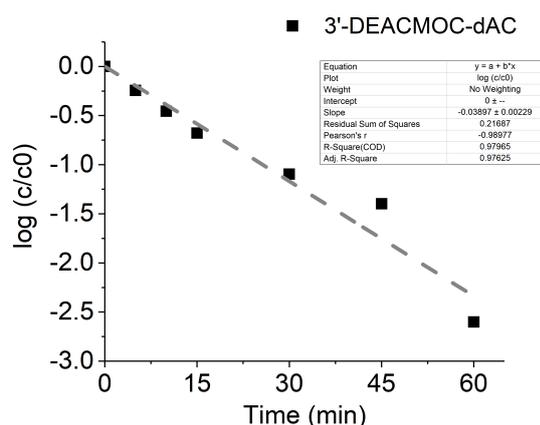
(B) mono-DMNPEOC-dAC, 64



(C) N-DEACMOC-dAC, 78



(D) 5'-NPEOC-dAC, 84

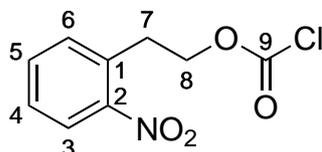


(E) 3'-NPEOC-dAC, 93

Fig. A.1 Linear Regression plots of Photouncaging Decays

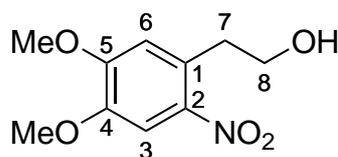
A.2 Characterisations

2-nitrophenethyl carbonochloridate (NPEOC-Cl) **58**¹⁸⁷



To a solution of 2-nitrophenethyl alcohol **57** (1.0 mL, 7.1 mmol) in dry tetrahydrofuran (THF) (20 mL) was added a solution of 20% *w/w* phosgene in dry toluene (7.5 mL, 14.18 mmol) dropwise at 0 °C under inert atmosphere. The reaction mixture was stirred at 0 °C for 1 h, followed by room temperature overnight. TLC analysis (5% MeOH/CH₂Cl₂) confirmed the complete conversion of the alcohol ($R_f = 0.48$) to the chloroformate ($R_f = 0.80$). Excess phosgene was removed under reduced pressure and was quenched with 0.1 M NaOH. 2-nitrophenethyl carbonochloridate **58** was isolated (1.53 g, 94%) as a viscous yellow oil. The crude product was used in the next step without further purification. ¹H NMR (600 MHz, CDCl₃) δ 7.96 (dd, 1 H, $J=7.6$, 1.0 Hz, *H*-3), 7.58 (td, 1 H, $J=7.6$, 1.0 Hz, *H*-5), 7.43 (td, 1 H, $J=7.6$, 1.2 Hz, *H*-4), 7.39 (dd, 1 H, $J=7.6$, 1.2 Hz, *H*-6), 4.61 (t, 2 H, $J=6.5$ Hz, *H*-8), 3.31 (t, 2 H, $J=6.5$ Hz, *H*-7); ¹³C NMR (150 MHz, CDCl₃) δ 150.55 (C-9), 149.33 (C-2), 133.63 (C-5), 133.08 (C-6), 131.67 (C-1), 128.60 (C-4), 125.32 (C-3), 71.19 (C-8), 32.40 (C-7); HRMS (CI) calcd for C₉H₈ClNO₄ [M]⁺ 229.01419, found 229.00250.

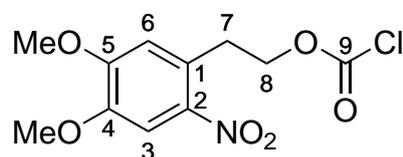
2-(4,5-dimethoxy-2-nitrophenyl)ethan-1-ol (DMNPE-OH) **62**²⁶⁶



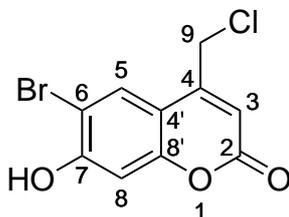
To a solution of 2-(4,5-dimethoxy-2-nitrophenyl)acetic acid **61** (519 mg, 2.15 mmol) in dry THF (13 mL) was added 1 M BH₃ · THF (5 mL, 5 mmol) under inert atmosphere. The solution was allowed to stir at room temperature for 2 h, then H₂O (2 mL) was slowly added to quench the reaction. The crude mixture was concentrated *in vacuo* and the resultant residue was diluted in ethyl acetate (100 mL) and washed with saturated NaHCO₃ (2 x 60 mL) and brine (60 mL), dried over MgSO₄ and concentrated *in vacuo* to afford 2-(4,5-dimethoxy-2-nitrophenyl)ethan-1-ol **62** (476 mg, 98%) as a

yellow precipitate: R_f 0.62 (5% MeOH/CH₂Cl₂); m.p.: 97 °C; ¹H NMR (600 MHz, DMSO-d₆) δ 7.54 (s, 1 H, *H*-3), 7.03 (s, 1 H, *H*-6), 4.74 (t, 1 H, $J=5.3$ Hz, -OH), 3.87 (s, 3 H, -OMe), 3.82 (s, 3 H, -OMe), 3.62 (q, 2 H, $J=6.8$ Hz, *H*-8), 3.02 (t, 2 H, $J=6.8$ Hz, *H*-7); ¹³C NMR (150 MHz, DMSO-d₆) δ 152.49 (C-4), 146.89 (C-5), 141.32 (C-2), 128.86 (C-1), 114.52 (C-6), 107.88 (C-3), 61.11 (C-8), 56.13 (MeO), 55.95 (MeO), 36.01 (C-7); IR v_{max} (solid): 3507 (O-H), 3322 (C-H), 2945 (C-H Ar), 1581 (C=C Ar), 1581-1325 (N-O) cm⁻¹; HRMS (CI) calcd for C₁₀H₁₃NO₅ [M]⁺ 227.07882, found 227.07888.

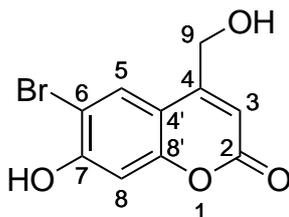
4,5-dimethoxy-2-nitrophenethyl carbonochloridate (DMNPEOC-Cl) **63**



To a solution of 2-(4,5-dimethoxy-2-nitrophenyl)ethan-1-ol **62** (450 mg, 1.98 mmol) in dry THF (10 mL) was added a solution of 20% *w/w* phosgene in dry toluene (2.5 mL, 4.73 mmol) dropwise at 0 °C under inert atmosphere. The reaction mixture was stirred at 0 °C for 1 h, followed by room temperature overnight. Excess phosgene was removed under reduced pressure and was quenched with 0.1 M NaOH. 4,5-dimethoxy-2-nitrophenethyl carbonochloridate **63** was isolated (550 mg, 96%) as dark yellow precipitate: R_f 0.94 (5% MeOH/CH₂Cl₂); m.p.: 64 °C. ¹H NMR (600 MHz, CDCl₃) δ 7.66 (s, 1 H, *H*-3), 6.73 (s, 1 H, *H*-6), 4.65 (t, 2 H, $J=6.3$ Hz, *H*-8), 3.98 (s, 3 H, -OMe), 3.95 (s, 3 H, -OMe), 3.35 (t, 2 H, $J=6.3$ Hz, *H*-7); ¹³C NMR (150 MHz, CDCl₃) δ 153.25 (C-4), 150.60 (C-9), 148.20 (C-5), 141.34 (C-2), 126.74 (C-1), 114.32 (C-6), 108.52 (C-3), 71.36 (C-8), 56.62 (MeO), 56.49 (MeO), 33.22 (C-7); HRMS (CI) calcd for C₁₁H₁₃ClNO₆ [M]⁺ 290.04259, found 290.04190.

6-Bromo-4-(chloromethyl)-7-hydroxymethylcoumarin (BHC-Cl) 70²¹⁷

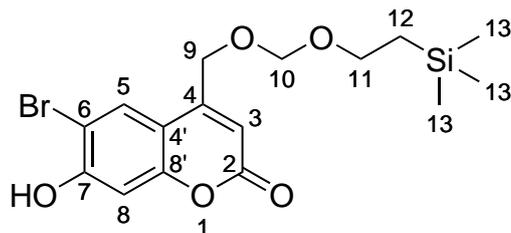
A solution of 4-bromoresorcinol **69** (1.00 g, 5.29 mmol) in methanesulfonic acid (8 mL) was treated with methyl 4-chloroacetoacetate (0.92 mL, 7.94 mmol) and was stirred for 2 h at room temperature. The resulting mixture was slowly poured into ice-water (24 mL) and then stirred for 30 min to give a white precipitate. The precipitate was collected by filtration and washed with cold water (3 x 5 mL). The resulting off-white precipitate was recrystallised with hexane/ethyl acetate (15 mL) to give compound **70** (0.96 g, 63%): R_f 0.17 (hexane/ethyl acetate 3:1); ^1H NMR (600 MHz, DMSO- d_6) δ 11.56 (s, 1 H, 7-OH), 7.99 (s, 1 H, H-5), 6.91 (s, 1 H, H-8), 6.47 (s, 1 H, H-3), 5.00 (d, 2 H, $J=0.9$ Hz, H-9); ^{13}C NMR (150 MHz, DMSO- d_6) δ 159.70 (C-2), 157.53 (C-7), 154.08 (C-6), 150.18 (C-4'), 129.09 (C-5), 112.14 (C-3), 110.70 (C-4), 106.19 (C-8'), 103.28 (C-8), 41.24 (C-9); HRMS (CI) calcd for $\text{C}_{10}\text{H}_6\text{BrClO}_3$ $[\text{M}+\text{H}]^+$ 288.92671, found 288.92624.

6-Bromo-7-hydroxy-4-hydroxymethylcoumarin (BHC-OH) 71²¹⁷

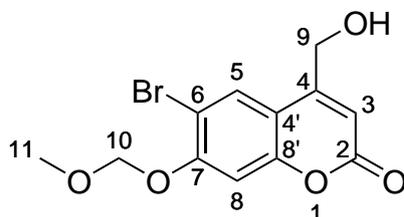
A suspension of 6-bromo-4-(chloromethyl)-7-hydroxymethylcoumarin **70** (0.92 g, 3.18 mmol) in water (230 mL) was heated at reflux for 5 days. The resulting mixture was cooled to room temperature concentrated *in vacuo*. The residue was triturated with hexane/diethyl ether (2:1, 6 mL) to afford 6-bromo-7-hydroxy-4-hydroxymethylcoumarin **71** (0.73 g, 85%). R_f 0.80 (ethyl acetate); ^1H NMR (600 MHz, DMSO- d_6) δ 11.42 (br, 1 H, 7-OH), 7.83 (s, 1 H, H-5), 6.89 (s, 1 H, H-8), 6.26 (d, 1 H, $J=1.6$ Hz, H-3), 5.63 (br, 1 H, 9'-OH), 4.69 (s, 2 H, H-9); ^{13}C NMR (150 MHz, DMSO- d_6) δ 160.17 (C-2), 157.09 (C-7), 156.06 (C-6), 153.68 (C-4'), 128.20 (C-5), 110.98 (C-3), 107.63 (C-4),

106.00 (C-8'), 103.07 (C-8), 59.07 (C-9); HRMS (CI) calcd for C₁₀H₇BrO₄ [M+H]⁺ 270.96060, found 270.96005.

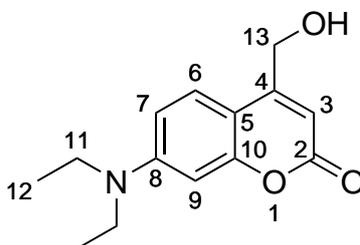
**6-bromo-7-*O*-(hydroxy)-4-((trimethylsilyl)ethoxymethyl)coumarin
(BHC-OSEM) 73**



To a solution of 6-bromo-7-hydroxy-4-hydroxymethylcoumarin **71** (150 mg, 0.55 mmol) in dry CH₂Cl₂ (12 mL) was added DIPEA (94 μL, 0.54 mmol) slowly and the solution was cooled to 0 °C. 2-(trimethylsilyl)ethoxymethyl chloride (96 μL, 0.54 mmol) was added dropwise. After being stirred at 0 °C for 2 h, the reaction mixture was diluted with CH₂Cl₂ and poured into 0.5 M citric acid (30 mL). The layers were separated and the organic layer was washed with citric acid (30 mL) followed by brine (30 mL). The residue was dried and concentrated *in vacuo* to afford a brown oil. Trituration with hexane afforded compound **73** (118 mg, 54%) as a light brown powder. R_f 0.69 (hexane/EtOAc 1:1); m.p.: 75 °C; ¹H NMR (600 MHz, DMSO-d₆) δ 11.45 (br, 1 H, 7-OH), 7.86 (s, 1 H, H-5), 6.90 (s, 1 H, H-8), 6.27 (s, 1H, H-3), 4.80 (s, 2 H, H-10), 3.36 (s, 2 H, H-9), 2.50 (s, 2 H, H-11), 0.84 (s, 2 H, H-12), 0.03 (s, 9 H, H-13); ¹³C NMR (150 MHz, DMSO-d₆) δ 160.08 (C-2), 153.91 (C-7), 153.71 (C-8'), 147.22 (C-4), 131.71 (C-5), 115.81 (C-4'), 111.71 (C-3), 111.41 (C-8), 110.54 (C-6), 95.51 (C-10), 69.12 (C-9), 66.65 (C-11), 24.01 (C-23), 2.12 (C-13); HRMS (CI) calcd for C₁₆H₂₁BrO₅Si [M+H]⁺ 401.33542, found 401.32444.

6-bromo-7-*O*-(methoxymethyl)-4-hydroxymethylcoumarin**(BHC-OMOM) 74**²¹⁷

A mixture of 6-bromo-7-hydroxy-4-hydroxymethylcoumarin **71** (200 mg, 0.74 mmol) and DIPEA (141 μ mL, 0.81 mmol) in dry CH₂Cl₂ (10 mL) was cooled to 0 °C, and chloromethyl methyl ether (62 μ L, 0.81 mmol) was added dropwise. After being stirred for 2 h at 0 °C, the reaction mixture was poured into 0.5 M citric acid (20 mL) and extracted with CHCl₃ (2 \times 20 mL). The combined organic layer was washed with brine (2 \times 30 mL) and dried over MgSO₄. After the solvent was removed *in vacuo*, the residue was triturated with hexane (2 \times 5 mL) to afford 6-bromo-7-*O*-(methoxymethyl)-4-hydroxymethylcoumarin **74** (132 mg, 57%): R_f0.34 (hexane/EtOAc 1:1); ¹H NMR (600 MHz, DMSO-d₆) δ 7.94 (s, 1 H, *H*-5), 7.26 (s, 1 H, *H*-8), 6.36 (s, 1 H, *H*-3), 5.68 (t, 1 H, *J*=5.5 Hz, -OH), 5.42 (s, 2 H, *H*-10), 4.72 (dd, 2 H, *J*=5.5, 1.7 Hz, *H*-9), 3.42 (s, 3 H, *H*-11); ¹³C NMR (150 MHz, DMSO-d₆) δ 159.98 (C-2), 157.32 (C-7), 155.15 (C-8'), 155.22 (C-4), 128.25 (C-5), 112.79 (C-4'), 108.96 (C-3), 106.14 (C-8), 103.54 (C-6), 94.81 (C-10), 59.09 (C-9), 56.28 (C-11); HRMS (CI) calcd for C₁₂H₁₂BrO₅ [M+H]⁺ 315.98633, found 315.98263.

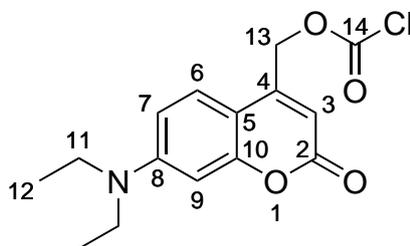
7-(diethylamino)-4-(hydroxymethyl)chroman-2-one (DEACM-OH) 76²¹⁸

To a solution of 4-diethylaminocoumarin **75** (4.63 g, 20 mmol) in hot *p*-xylene (120 mL) was added SeO₂ (3.33 g, 30 mmol) and the mixture was heated under reflux with vigorous stirring overnight. The mixture was then filtered and concentrated

under reduced pressure. The dark brown residual oil was dissolved in a mixture of ethanol and THF (1:1, 120 mL) and sodium borohydride (380 mg, 10 mmol) was added and the solution was stirred overnight at room temperature. The suspension was carefully hydrolysed with 1 M HCl (20 mL), diluted with water (30 mL) and extracted with CH₂Cl₂ (3 x 50 mL). The organic phase was washed with water (2 x 100 mL), brine (100 mL) and dried over MgSO₄ and concentrated *in vacuo*. The crude material was purified by flash chromatography (CH₂Cl₂/Acetone 5:1) to yield 7-(diethylamino)-4-(hydroxymethyl)chroman-2-one **76** as a yellow solid (2.71 g, 55%); ¹H NMR (600 MHz, DMSO-d₆) δ 7.39 (d, 1 H, *J*=9.0 Hz *H*-6), 6.62 (dd, 1 H, *J*=9.0, 2.7 Hz, *H*-7), 6.48 (d, 1 H, *J*=2.7 Hz, *H*-9), 6.03 (s, 1 H, *H*-3), 5.51 (t, 1 H, *J*=5.6 Hz, -OH), 4.63 (d, 2 H, *J*=5.6 Hz, *H*-13), 3.38 (q, 4 H, *J*=7.2 Hz, *H*-11), 1.0 (t, 6 H, *J*=7 Hz, *H*-12); ¹³C NMR (150 MHz, DMSO-d₆) δ 161.21(C-2), 156.97 (C-10), 155.64 (C-8), 150.16 (C-4), 125.11 (C-6), 108.53 (C-3), 105.64 (C-5), 103.80 (C-7), 96.73 (C-9), 59.05 (C-13), 43.99 (C-11), 12.17 (C-12); HRMS (CI) calcd for C₁₄H₁₇NO₃ [M]⁺ 247.12084, found 247.120289.

(7-(diethylamino)-2-oxochroman-4-yl)methyl carbonochloridate

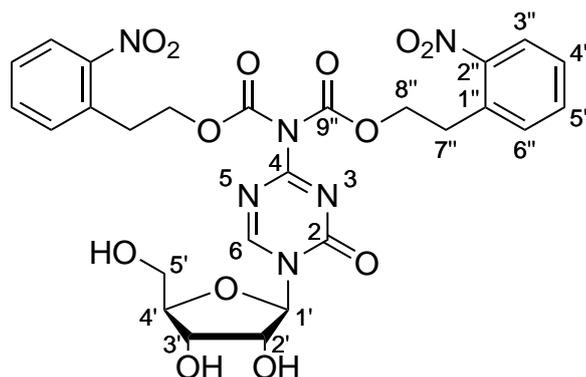
(DEACMOC-Cl) **77**²¹⁹



A solution of 7-(diethylamino)-4-(hydroxymethyl)chroman-2-one **76** (550 mg, 2.23 mmol) in dry THF (18 mL) was treated with DIPEA (522 μL, 3.34 mmol) and stirred for 5 min at 0 °C. After the dropwise addition of 15% w/w phosgene (2.94 mL, 5.58 mmol), stirring was continued under the exclusion of light at 0 °C for 3 h. The reaction mixture was transferred onto a mixture of ethyl acetate and water (1:1, 250 mL) and the layers were separated and the organic layer was dried over MgSO₄ and concentrated *in vacuo*. The crude chloroformate **77** was isolated as a yellow solid (656 mg, 95%) and was used in the next step without further purification. ¹H NMR (600 MHz, DMSO-d₆) δ 7.41 (d, 1 H, *J*=9.0 Hz *H*-6), 6.62 (dd, 1 H, *J*=9.0, 2.7 Hz, *H*-7), 6.52 (d, 1 H, *J*=2.7 Hz, *H*-9), 6.01 (s, 1 H, *H*-3), 4.65 (d, 2 H, *J*=5.6 Hz, *H*-13), 3.41 (q, 4 H, *J*=7.2 Hz, *H*-11), 1.0 (t, 6 H, *J*=7 Hz, *H*-12); ¹³C NMR (150 MHz, DMSO-d₆)

δ 161.21(C-2), 156.97 (C-10), 155.64 (C-8), 150.16 (C-4), 148.23 (C-14) 125.11 (C-6), 108.53 (C-3), 105.64 (C-5), 103.80 (C-7), 96.73 (C-9), 59.05 (C-13), 43.99 (C-11), 12.17 (C-12); HRMS (CI) calcd for $C_{15}H_{16}ClNO_4$ $[M]^+$ 310.52084, found 310.520289.

4-*N*-[di-(2-nitrophenyl)ethoxycarbonyl]-5-azacytidine (Bis-NPEOC-AC) **59**

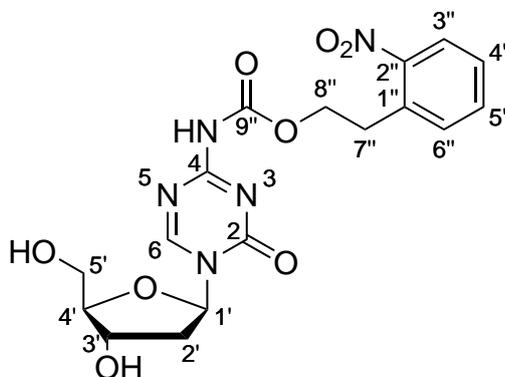


5-azacytidine **5** (150 mg, 0.61 mmol) was coevaporated with anhydrous pyridine (3 x 5 mL) and subsequently suspended into anhydrous DMF (10 mL) before adding hexamethyldisilazane (650 μ L, 3.11 mmol). After being stirred for 2 h at room temperature, the solution was concentrated, and the residue was dried by three cycles of evaporation from toluene (5 mL), followed by anhydrous pyridine (3 x 5 mL). The resultant residue was subsequently suspended into anhydrous pyridine (30 mL) to which chloroformate **58** (1.53 g, 6.69 mmol) in anhydrous CH_2Cl_2 (8 mL) was added and the resultant mixture was stirred at room temperature overnight. The solvent was removed under reduced pressure and the resultant residue was coevaporated with toluene (3 x 10 mL) and suspended in anhydrous DMF (10 mL). Silyl deprotection was achieved with the addition of tris(dimethylamino) sulfonium difluorotrimethylsilicate (TAS-F)(600 mg, 2.17 mmol). After 2 h, the mixture was concentrated *in vacuo* and the resultant residue was coevaporated with anhydrous toluene (3 x 5 mL). The crude material was purified by flash chromatography (0-10% EtOH/DCM, 10% EtOH/DCM R_f 0.58) to yield bis-NPEOC-AC **59** as a colourless film (56.9 mg, 15%). 1H NMR (600 MHz, $DMSO-d_6$) δ 9.10 (s, 1 H, $H-6$), 7.93 (d, 2 H, $J=8.2$ Hz, $H-3''$), 7.64 (td, 2 H, $J=7.6, 1.2$ Hz, $H-5''$), 7.51-7.48 (m, 4 H, $H-4''$, $H-6''$), 5.74 (d, 1 H, $J=4.8$ Hz, $3'-OH$), 5.68 (d, 1 H, $J=1.4$ Hz, $H-1'$), 5.34 (t, 1 H, $J=4.6$ Hz, $5'-OH$), 5.11 (d, 1 H, $J=6.8$ Hz, $2'-OH$), 4.46 (tt, 4 H, $J=6.0, 2.9$ Hz, $H-8''$), 4.14 (dt, 1 H, $J=4.7, 2.4$ Hz, $H-2''$), 4.09 (dt, 1 H, $J=7.0, 3.5$ Hz, $H-3'$), 3.97 (dt, 1 H, $J=7.8, 2.2$ Hz, $H-4'$), 3.85 (ddd, 1 H, $J=12.0, 4.5, 2.6$ Hz, $H-5'a$), 3.64 (ddd, $J=12.1, 4.4, 2.1$ Hz, $H-5'b$), 3.18 (t, 4 H,

$J=6.2$ Hz, $H-7''$); ^{13}C NMR (150 MHz, DMSO-d_6) δ 163.59 (C-4), 159.03 (C-6), 152.66 (C-9''), 150.16 (C-2), 149.31 (C-2''), 133.36 (C-5''), 132.67 (C-6''), 131.78 (C-1''), 128.134 (C-3''), 118.14 (C-1'), 91.24 (C-1'), 83.97 (C-4'), 74.02 (C-2'), 67.66 (C-8''), 67.50 (C-3'), 58.86 (C-5'), 30.62 (C-7''); IR ν_{max} (oil): 3292 (O-H), 1755 (C=O NPEOC), 1520 (N-O) cm^{-1} ; HRMS (ESI) calcd for $\text{C}_{26}\text{H}_{26}\text{N}_6\text{O}_{13}\text{Na}$ $[\text{M}+\text{Na}]^+$ 653.1456, found 653.1456.

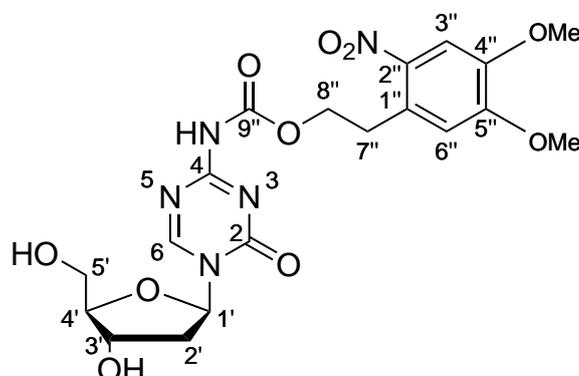
General synthesis of mono-N-tagged dAC and dC derivatives

5-aza-2'-deoxycytidine **6** (0.66 mmol) or 2'-deoxycytidine **7** (1 mmol) was dried after three successive evaporations of a solution in anhydrous pyridine (5 mL). The residue was dissolved in anhydrous DMF (10 mL) and treated with hexamethyldisilazane (2.5 eq). After 2 h of stirring at room temperature, the solution was concentrated and the residue was dried by three cycles of evaporation from toluene (5 mL) followed by anhydrous pyridine (3 x 5 mL) before finally being dissolved in anhydrous pyridine (20 mL). To this, a solution of the relevant chloroformate (2 eq) in dry CH_2Cl_2 (8 mL) was added and the mixture was left stirring overnight. The solution was concentrated to dryness and traces of pyridine were removed by coevaporation with toluene (3 x 10 mL). The resultant residue was suspended in dry DMF (10 mL) and TAS-F (1.5 eq) was added to achieve silyl deprotection. After 2 h, the mixture was concentrated *in vacuo* and the resultant residue was coevaporated with anhydrous toluene (3 x 5 mL). The crude material was purified by flash chromatography (0-10% MeOH/DCM) to yield the title compounds.

2'-Deoxy-4-N-[mono-(2-nitrophenyl)ethoxycarbonyl]-5-azacytidine**(mono-NPEOC-dAC) 60**

Isolated as a white foam (231 mg, 53%); 5% MeOH/DCM R_f 0.16; m.p. 46 °C to 48 °C; ^1H NMR (600 MHz, DMSO- d_6) δ 10.71 (br, 1 H, NH), 8.81 (s, 1 H, H-6), 7.97 (dd, 1 H, $J=8.1, 1.1$ Hz, H-3''), 7.67 (td, 1 H, $J=7.6, 1.2$ Hz, H-5''), 7.60 (dd, 1 H, $J=7.7, 1.1$ Hz, H-6''), 7.54-7.50 (m, 1 H, H-4''), 5.98 (t, 1 H, $J=5.1$ Hz, H-1'), 5.36 (d, 1 H, $J=4.4$ Hz, 3'-OH), 5.19 (t, 1 H, $J=5.1$ Hz, 5'-OH), 4.36 (t, 2 H, $J=6.5$ Hz, H-8''), 4.25-4.22 (m, 1 H, H-3'), 3.87 (q, 1 H, $J=3.7$ Hz, H-4'), 3.64 (ddd, 1 H, $J=12.0, 5.0, 3.5$ Hz, H-5'a), 3.57 (ddd, 1 H, $J=12.0, 5.0, 3.5$ Hz, H-5'b), 3.19-3.15 (m, 2 H, H-7''), 2.30 (ddd, 1 H, $J=13.3, 6.3, 4.8$ Hz, H-2'a), 2.20 (dt, 1 H, $J=13.4, 5.9$ Hz, H-2'b); ^{13}C NMR (150 MHz, DMSO- d_6) δ 162.46 (C-4), 157.33 (C-6), 153.07 (C-9''), 150.73 (C-2), 149.32 (C-2''), 133.42 (C-5''), 132.95 (C-6''), 132.13 (C-1''), 128.25 (C-4''), 124.53 (C-3''), 88.02 (C-4'), 86.36 (C-1'), 69.31 (C-3'), 64.38 (C-8''), 60.42 (C-5'), 40.63 (C-2'), 31.41 (C-7''); IR ν_{max} (solid): 3291 (O-H), 2929 (N-H), 1756 (C=O NPEOC), 1681 (C=O Aza), 1520 (N-O), 1435 (N-O), 1180 (C-O), 1050 (C-O) cm^{-1} ; HRMS (ESI) calcd for $\text{C}_{17}\text{H}_{19}\text{N}_5\text{O}_8$ $[\text{M}]^-$ 420.1155, found 420.1135.

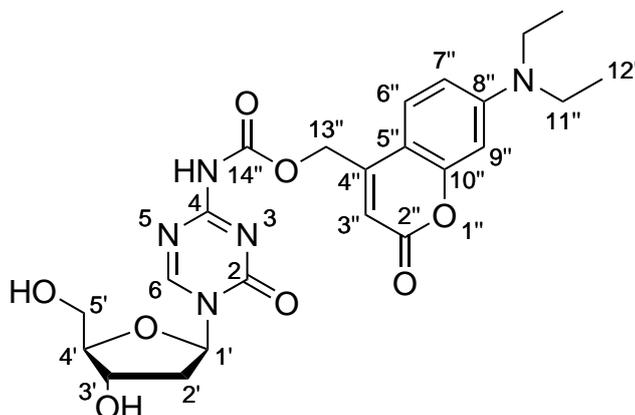
2'-Deoxy-4-*N*-[mono-[2-(4,5-dimethoxy-2-nitrophenyl)]ethoxycarbonyl]-5-azacytidine (mono-DMNPEOC-dAC) 64



Isolated as a yellow foam (140 mg, 28%); 5% MeOH/DCM R_f 0.39; m.p. 106 °C to 108 °C. ^1H NMR (600 MHz, DMSO- d_6) δ 10.73 (br, 1 H, NH), 8.82 (s, 1 H, *H*-6), 7.61 (s, 1 H, *H*-3''), 7.05 (s, 1 H, *H*-6''), 5.99 (t, 1 H, $J=6.0$ Hz, *H*-1'), 5.28 (d, 1 H, $J=4.4$ Hz, 3'-OH), 5.13 (t, 1 H, $J=5.1$ Hz, 5'-OH), 4.37 (t, 2 H, $J=6.5$ Hz, *H*-8''), 4.25-4.22 (m, 1 H, *H*-3'), 3.89 (s, 3 H, $-\text{OCH}_3$), 3.88-3.87 (m, 1 H, *H*-4'), 3.84 (s, 3 H, $-\text{OCH}_3$), 3.67-3.63 (m, 1 H, *H*-5'a), 3.59-3.53 (m, 1 H, *H*-5'b), 3.20 (t, 2 H, $J=6.5$ Hz, *H*-7''), 2.30 (ddd, 1 H, $J=13.3, 6.3, 4.8$ Hz, *H*-2'a), 2.21 (dt, 1 H, $J=13.4, 5.9$, Hz, *H*-2'b); ^{13}C NMR (150 MHz, DMSO- d_6) δ 163.46 (C-4), 157.35 (C-6), 152.74 (C-5''), 150.69 (C-9''), 147.28 (C-4''), 141.00 (C-2''), 127.37 (C-1''), 114.75 (C-6''), 108.06 (C-3''), 88.07 (C-5'), 86.33 (C-1'), 69.32 (C-3'), 64.38 (C-8''), 60.44 (C-5'), 56.25 ($-\text{OMe}$), 55.97 ($-\text{OMe}$), 40.64 (C-2'), 32.20 (C-7''); IR ν_{max} (solid): 3351 (O-H), 2944 (N-H), 1743 (C=O DMNPEOC), 1680 (C=O Aza), 1518 (N-O), 1461 (N-O), 1182 (C-O), 1060 (C-O) cm^{-1} ; HRMS (ESI) calcd for $\text{C}_{19}\text{H}_{22}\text{N}_5\text{O}_{10}\text{Na}$ $[\text{M}+\text{Na}]^+$ 480.1367, found 480.1369.

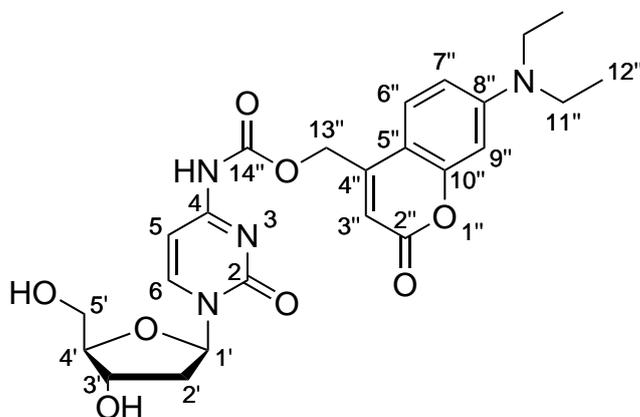
2'-Deoxy-4-*N*-[mono-7-(diethylamino)-2-oxochroman-4-yl]methoxycarbonyl]-

5-azacytidine (N-DEACMOC-dAC) 78

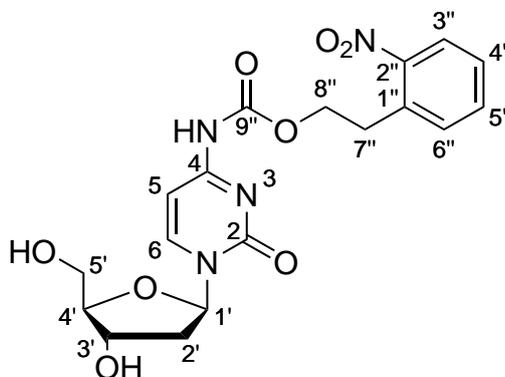


Isolated as a yellow solid (52 mg, 16%); 5% MeOH/DCM R_f 0.45); m.p. 148 °C to 150 °C. ^1H NMR (600 MHz, DMSO- d_6) δ 11.04 (br, 1 H, NH), 8.86 (s, 1 H, H-6), 7.47 (d, 1 H, $J=9.0$ Hz, H-6''), 6.70 (dd, 1 H, $J=9.2, 2.4$ Hz, H-7''), 6.55 (d, 1 H, $J=2.4$ Hz, H-9''), 6.17 (s, 1 H, H-3''), 6.00 (t, 1 H, $J=6.3$ Hz, H-1'), 5.37 (s, 2 H, H-13''), 5.29 (d, 1 H, $J=4.4$ Hz, 3'-OH), 5.14 (t, 1 H, $J=5.2$ Hz, 5'-OH), 4.25 (m, 1 H, H-3'), 3.87 (q, 1 H, $J=3.8$ Hz, H-4'), 3.65 (ddd, 1 H, $J=11.7, 5.2, 3.6$ Hz, H-5'a), 3.57 (ddd, 1 H, $J=11.7, 5.2, 3.6$ Hz, H-5'b), 3.43 (q, 4 H, $J=7.2$ Hz, H-11''), 2.31 (ddd, 1 H, $J=13.6, 6.5, 4.7$ Hz, H-2'a), 2.22 (dt, 1 H, $J=13.4, 5.9$ Hz, H-2'b), 1.12 (t, 6 H, $J=7.0$ Hz, H-12''); ^{13}C NMR (150 MHz, DMSO- d_6) δ 160.78 (C-2), 160.73 (C-2''), 157.56 (C-10''), 155.76 (C-8''), 153.04 (C-4), 150.76 (C-14''), 150.48 (C-6), 125.45 (C-6''), 108.77 (C-3''), 105.07 (C-5''), 104.88 (C-7''), 96.86 (C-9''), 88.13 (C-4'), 86.43 (C-1'), 69.32 (C-3'), 61.99 (C-13''), 60.45 (C-5'), 44.05 (C-11''), 40.68 (C-2'), 12.32 (C-12''); IR v_{max} (solid): 3239 (O-H), 2942 (N-H), 1680 (C=O), 1072 (C-O) cm^{-1} ; HRMS (ESI) calcd for $\text{C}_{23}\text{H}_{27}\text{N}_5\text{O}_8\text{Na}$ $[\text{M}+\text{Na}]^+$ 524.1757, found 524.1755.

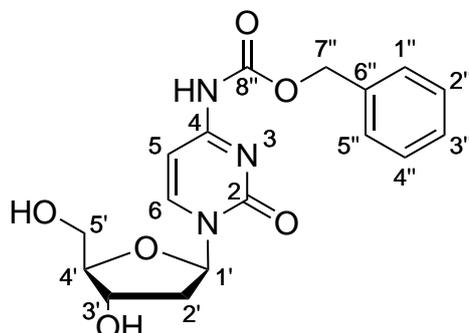
2'-Deoxy-4-*N*-[mono-7-(diethylamino)-2-oxochroman-4yl]methoxycarbonyl
cytidine (N-DEACMOC-dC) 81



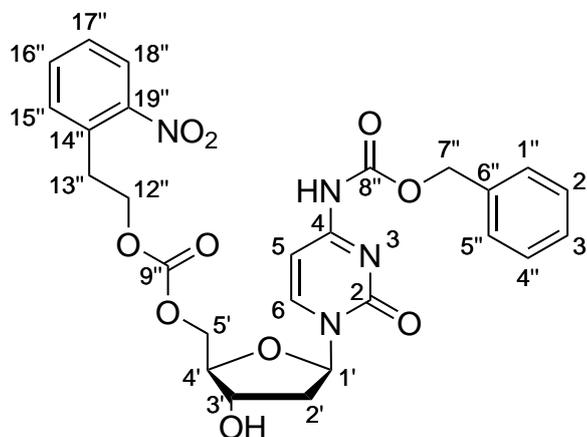
Isolated as a yellow solid (400.1 mg, 80%); 5% MeOH/DCM R_f 0.48); m.p. 155 °C to 156 °C. ^1H NMR (600 MHz, DMSO- d_6) δ 11.07 (br, 1 H, NH), 8.35 (d, 1 H, $J=7.5$ Hz, $H-5$), 7.47 (d, 1 H, $J=9.0$ Hz, $H-6''$), 7.04 (d, 1 H, $J=7.4$ Hz, $H-6$), 6.70 (dd, 1 H, $J=9.1, 2.3$ Hz, $H-7''$), 6.55 (d, 1 H, $J=1.7$ Hz, $H-9''$), 6.15 (s, 1 H, $H-3''$), 6.11 (dd, 1 H, $J=6.5, 6.2$ Hz, $H-1'$), 5.39 (s, 2 H, $H-13''$), 5.28 (d, 1 H, $J=4.2$ Hz, $3'-OH$), 5.06 (dd, 1 H, $J=5.5, 5.0$ Hz, $5'-OH$), 4.22-4.20 (m, 1 H, $H-3'$), 3.85 (ddd, 1 H, $J=3.9, 3.6, 3.3$ Hz, $H-4'$), 3.63-3.60 (m, 1 H, $H-5'a$), 3.56 (ddd, 1 H, $J=11.8, 4.7, 4.5$ Hz, $H-5'b$), 3.43 (q, 4 H, $J=6.8$ Hz, $H-11''$), 2.28 (ddd, 1 H, $J=13.2, 6.0, 3.8$ Hz, $H-2'a$), 2.02 (dt, 1 H, $J=13.4, 6.3$ Hz, $H-2'b$), 1.12 (t, 6 H, $J=7.1$ Hz, $H-12''$); ^{13}C NMR (150 MHz, DMSO- d_6) δ 160.76 (C-2), 160.76 (C-2''), 155.77 (C-8''), 154.32 (C-4), 153.21 (C-14''), 152.58 (C-10''), 150.49 (C-6), 145.00 (C-5), 125.41 (C-6''), 108.77 (C-7''), 105.07 (C-3''), 104.88 (C-8''), 96.87 (C-9''), 94.11 (C-6), 87.91 (C-4'), 86.16 (C-1'), 69.95 (C-3'), 62.19 (C-13''), 60.95 (C-5'), 44.05 (C-11''), 40.16 (C-2'), 12.37 (C-12''); IR ν_{max} (solid): 3191 (O-H), 2926 (N-H), 1740 (C=O), 1642 (C=O Aza), 1604 (C=C), 1272 (C-N), 1063 (C-O) cm^{-1} ; HRMS (ESI) calcd for $\text{C}_{24}\text{H}_{29}\text{N}_4\text{O}_8$ $[\text{M}+\text{H}]^+$ 501.1985, found 501.1986.

2'-Deoxy-4-N-[mono-(2-nitrophenyl)ethoxycarbonyl]-cytidine**(N-NPEOC-dC) 80**

Isolated as a white foam (208 mg, 50%); 5% MeOH/DCM R_f 0.19; m.p. 57 °C to 58 °C. ^1H NMR (600 MHz, DMSO- d_6) δ 10.72 (br, 1 H, NH), 8.29 (d, 1 H, $J=7.5$ Hz, $H-5$), 7.98 (d, 1 H, $J=8.2$ Hz, $H-3''$), 7.68 (t, 1 H, $J=7.4$ Hz, $H-5''$), 7.60 (dd, 1 H, $J=7.7$ Hz, $H-6''$), 7.52 (t, 1 H, $J=7.8$ Hz, $H-4''$), 6.98 (d, 1 H, $J=7.5$ Hz, $H-6$), 6.09 (dd, 1 H, $J=6.8, 6.1$ Hz, $H-1'$), 5.27 (d, 1 H, $J=4.2$ Hz, $3'-\text{OH}$), 5.06 (t, 1 H, $J=5.2$ Hz, $5'-\text{OH}$), 4.39 (t, 2 H, $J=6.5$ Hz, $H-8''$), 4.21-4.19 (m, 1 H, $H-3'$), 3.84 (q, 1 H, $J=3.6$ Hz, $H-4'$), 3.61-3.59 (m, 1 H, $H-5'a$), 3.57-3.54 (m, 1 H, $H-5'b$), 3.18 (t, 2 H, $J=6.6$ Hz, $H-7''b$), 2.28-2.24 (m, 2 H, $H-2''a$), 2.00 (ddd, 1 H, $J=13.5, 6.8, 5.7$ Hz, $H-2'b$); ^{13}C NMR (150 MHz, DMSO- d_6) δ 162.70 (C-4), 155.28 (C-2), 153.03 (C-9''), 149.37 (C-4), 144.69 (C-5), 133.41 (C-5''), 133.02 (C-6''), 132.13 (C-2''), 128.26 (C-4''), 124.56 (C-3''), 94.17 (C-6), 87.87 (C-4'), 86.05 (C-1'), 69.95 (C-3'), 64.54 (C-8''), 60.97 (C-5'), 40.85 (C-2'), 18.61 (C-7''), 157.33 (C-6), 153.07 (C-9''), 149.32 (C-2''), 133.42 (C-5''), 132.95 (C-6''), 132.13 (C-1''), 128.25 (C-4''), 124.53 (C-3''), 88.02 (C-4'), 86.36 (C-1'), 69.31 (C-3'), 64.38 (C-8''), 60.42 (C-5'), 40.63 (C-2'), 31.41 (C-7''); IR v_{max} (solid): 3273 (O-H), 2931 (N-H), 1742 (C=O NPEOC), 1642 (C=O Aza), 1494 (N-O), 1342 (N-O), 1187 (C-O), 1054 (C-O) cm^{-1} ; HRMS (ESI) calcd for $\text{C}_{18}\text{H}_{20}\text{N}_4\text{O}_8$ $[\text{M}+\text{H}]^+$ 421.1359, found 421.1349

2'-Deoxy-4-*N*-[mono-carboxybenzoyl] cytidine (N-Cbz-dC) 85²⁶⁷

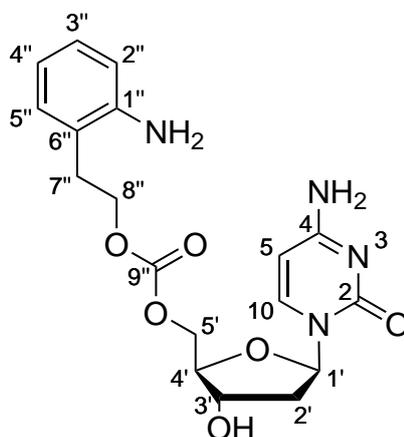
Isolated as a pink solid (287 mg, 79%); ¹H NMR (600 MHz, DMSO-d₆) δ 10.83 (br, 1 H, *NH*), 8.31 (d, 1 H, *J*=7.4 Hz, *H*-5), 7.42-7.38 (m, 4 H, *H*-1'', *H*-2'', *H*-5'', *H*-7''), 7.36-7.33 (m, 1 H, *H*-3''), 7.04 (d, 1 H, *J*=7.3 Hz, *H*-5), 6.11 (t, 1 H, *J*=6.3 Hz, *H*-1'), 5.27 (d, 1 H, *J*=4.2 Hz, 3'-*OH*), 5.18 (s, 2 H, *H*-7''), 5.05 (t, 1 H, *J*=5.3 Hz, 5'-*OH*), 4.22-4.20 (m, 1 H, *H*-3'), 3.85 (q, 1 H, *J*=3.7 Hz, *H*-4'), 3.61(ddd, 1 H, *J*=11.9, 5.5, 3.8 Hz, *H*-5'a), 3.56 (ddd, 1 H, *J*=11.9, 5.3, 4.0 Hz *H*-5'b), 2.27 (ddd, 1 H, *J*=13.4, 6.2, 3.9 Hz, *H*-2'a), 2.01 (ddd, 1 H, *J*=13.4, 6.7, 5.9 Hz, *H*-2') b; ¹³C NMR (150 MHz, DMSO-d₆) δ 162.76 (C-2), 154.30 (C-4), 153.15 (C-8''), 144.76 (C-6), 135.97 (C-6''), 128.52 (C-1'', C-5''), 128.19 (C-2'', C-5''), 127.97 (C-3''), 94.18 (C-5), 87.88 (C-4'), 86.08 (C-1'), 69.95 (C-3'), 66.52 (C-7''), 60.97 (C-5''), 40.87 (C-2'); HRMS (ESI) calcd for C₁₇H₂₀N₃O₆ [M+H]⁺ 362.1352, found 362.1339.

2'-Deoxy-5'-(2-nitrophenyl)ethoxycarbonyl-4-*N*-[mono-carboxybenzoyl] cytidine (N-Cbz-5'-NPEOC-dC) 86

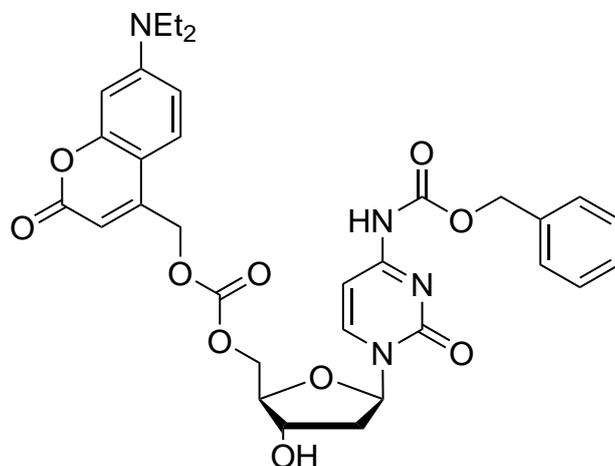
Analogue **85** (100 mg, 0.28 mmol) was dried after three successive evaporations of a solution in anhydrous pyridine (3 mL). The residue was dissolved in anhydrous pyridine (3 mL) and was treated with chloroformate **58** (120 mg, 0.52 mmol) in anhydrous CH_2Cl_2 (3 mL). After 5 h of stirring at room temperature, the solution was concentrated and the residue was coevaporated with anhydrous toluene (3 x 3 mL). The residue was diluted with water (15 mL) and the aqueous layer was back-extracted with CH_2Cl_2 (3x 50 mL). The organic layers were combined, dried and concentrated. The crude was purified via flash column chromatography (0-10% MeOH/DCM) to yield the title compound **86** (84 mg, 55%) as a white solid; m.p. 51 °C to 53 °C ^1H NMR (600 MHz, DMSO- d_6) δ 10.86 (br, 1 H, NH), 8.01 (d, 1 H, $J=7.6$ Hz, $H-6$), 7.95 (dd, 1 H, $J=7.9$, 1.4 Hz, $H-18''$), 7.66 (td, 1 H, $J=7.6$, 1.4 Hz, $H-16''$), 7.54 (dd, 1 H, $J=7.7$, 1.5 Hz, $H-15''$), 7.49 (td, 1 H, $J=7.7$, 1.5 Hz, $H-17''$), 7.42-7.38 (m, 4 H, $H-1''$, $H-2''$, $H-5''$, $H-7''$), 7.36-7.33 (m, 1 H, $H-3''$), 7.04 (d, 1 H, $J=7.5$ Hz, $H-5$), 6.12 (t, 1 H, $J=6.4$ Hz, $H-1'$), 5.74 (d, 1 H, $J=4.4$ Hz, 3'-OH), 5.20 (s, 2 H, $H-7''$), 4.37 (t, 2 H, $J=6.4$ Hz, $H-12''$), 4.29 (dd, 1 H, $J=11.8$, 3.6 Hz, $H-5'a$), 4.24 (dd, 1 H, $J=11.8$, 5.9 Hz, $H-5'b$), 4.18 (dq, 1 H, $J=6.3$, 4.2 Hz, $H-3'$), 4.0 (dt, 1 H, $J=5.9$, 3.8 Hz, $H-4'$), 3.20 (t, 2 H, $J=6.4$ Hz, $H-13''$), 2.29 (ddd, 1 H, $J=13.6$, 6.3, 4.2 Hz, $H-2'a$), 2.07 (dt, 1 H, $J=13.6$, 6.6 Hz, $H-2'b$); ^{13}C NMR (150 MHz, DMSO- d_6) δ 162.86 (C-2), 154.19 (C-4), 153.15 (C-8''), 149.46 (C-9''), 144.54 (C-6), 135.96 (C-6''), 133.41 (C-15''), 132.66 (C-19''), 128.51 (C-1'', C-5''), 128.31 (C-16''), 128.20 (C-2'', C-5''), 127.98 (C-3''), 124.59 (C-18''), 94.49 (C-5), 86.23 (C-4'), 84.08 (C-1'), 70.04 (C-3'), 67.33.97 (C-5''), 67.18 (C-12''), 66.56 (C-7''), 48.63 (C-2'), 31.17 (C-13''); IR v_{max} (solid): 3297 (O-H), 2922 (N-H), 1738 (C=O), 1644 (C=O), 1494 (N-O), 1192 (C-O), 1091 (C-O) cm^{-1} ; HRMS

(ESI) calcd for $C_{26}H_{27}N_4O_{10}$ $[M+H]^+$ 555.1727, found 555.1718.

2'-Deoxy-5'-[(2-aminophenyl)ethoxycarbonyl]-cytidine (5-APEOC-dC) **87**



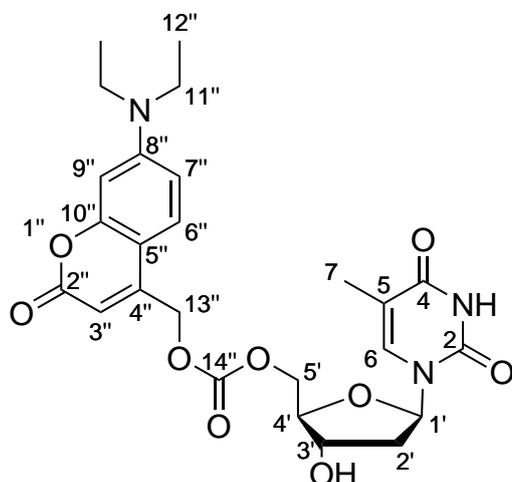
Palladium on activated carbon (10 wt%, 10 mg, 0.009 mmol) was added to a stirred solution of compound **86** (50 mg, 0.09 mmol) in anhydrous methanol (5 mL) under inert atmosphere. The reaction flask was evacuated and flushed with H_2 . The resulting mixture was stirred at room temperature for 1 h. After completion of the reaction, the catalyst was filtered through a pad of celite and the filter was washed with methanol (10 mL). The filtrate was concentrated under reduced pressure to obtain **87** as a white film (33 mg, 94%). 1H NMR (600 MHz, DMSO- d_6) δ 7.79 (d, 1 H, $J=7.4$ Hz, $H-5$), 7.13 (d, 1 H, $J=35.2$ Hz, dC-NH2), 6.93-6.86 (m, 2 H, $H-4''$, $H-5''$), 6.60 (dd, 1 H, $J=12.1$, 7.7 Hz, $H-3''$), 6.47 (q, 1 H, $J=7.1$ Hz, $H-6''$), 6.15 (t, 1 H, $J=6.5$ Hz, $H-1'$), 5.71 (d, 1 H, $J=7.4$ Hz, $H-5$), 4.95 (b, 1 H, -NH $_2$), 4.82 (s, 1 H, 3'-OH), 4.20 (t, 1 H, $J=7.1$ Hz, $H-3'$), 4.2-4.05 (m, 2 H, $H-5'$), 3.56- 3.52 (m, 1 H, $H-4'$), 2.78 (t, 2 H, $J=7.1$ Hz, $H-8''$), 2.58 (t, 2 H, $J=7.0$ Hz, $H-7''$), 2.11-2.08 (m, 1 H, $H2'a$), 1.96-1.90 (m, 1 H, $H2'b$); ^{13}C NMR (150 MHz, DMSO- d_6) δ 165.56 (C-4), 155.27 (C-2), 146.50 (C-2''), 141.00 (C-6), 130.06 (C-1''), 127.34 (C-4''), 126.65 (C-5''), 123.15 (C-9''), 116.17 (C-6''), 114.77 (C-3''), 93.94 (C-5), 87.21 (C-4'), 84.86 (C-1'), 70.43 (C-5'), 66.15 (C-3'), 40.38 (C-2'), 34.63 (C-8''), 30.06 (C-7''); HRMS (ESI) calcd for $C_{18}H_{22}N_4O_6$ $[M+H]^+$ 391.3960, found 391.3553.

2'-Deoxy-5'-[7-(diethylamino)-2-oxochroman-4-yl]methoxycarbonyl]-4-N-[monocarboxybenzoyl] cytidine (N-Cbz-5'-DEACMOC-dC) 88

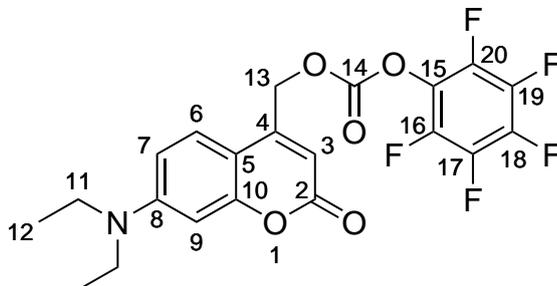
Analogue **85** (237 mg, 0.66 mmol) was dried after three successive evaporations of a solution in anhydrous pyridine (3 mL). The residue was dissolved in anhydrous pyridine (3 mL) and was treated with chloroformate **77** (600 mg, 1.94 mmol) in anhydrous CH_2Cl_2 (3 mL). After overnight stirring at room temperature, the solution was concentrated and the residue was coevaporated with anhydrous toluene (3 x 3 mL). The residue was diluted with water (15 mL) and the aqueous layer was back-extracted with CH_2Cl_2 (3x 50 mL). The organic layers were combined, dried and concentrated. The crude was purified via flash column chromatography (0-10% MeOH/DCM) to yield compound **88** (15 mg, 3.5%) as a yellow film. Identity of the product was confirmed *via* LC-MS. The analogue was not submitted for full NMR characterization in order to have enough material for the next hydrogenation step.

Hydrogenation of (N-Cbz-5'-DEACMOC-dC) 88

Palladium on activated carbon (10 wt%, 2.1 mg, 0.002 mmol) was added to a stirred solution of compound **88** (14.7 mg, 0.02 mmol) in anhydrous methanol (3 mL) under inert atmosphere. The reaction flask was evacuated and flushed with H_2 . The resulting mixture was stirred at room temperature for 1 h. No product was detected using this method even after 24 h.

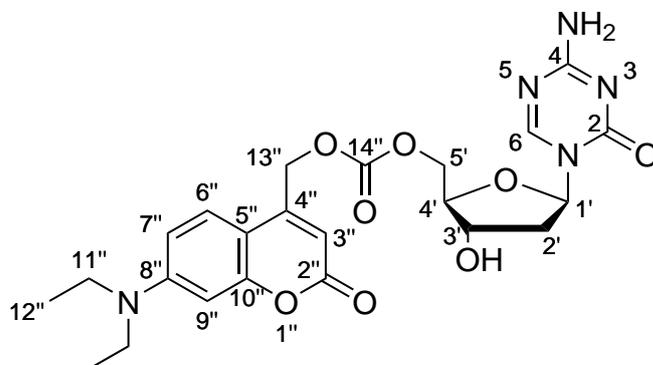
2'-Deoxy-5'-[7-(diethylamino)-2-oxochroman-4yl]methoxycarbonyl]-thymidine (5-DEACMOC-dT) **83**

Thymidine **82** (74 mg, 0.3 mmol) was dried after three successive evaporations of a solution in anhydrous pyridine (3 mL). The residue was dissolved in anhydrous pyridine (3 mL) and was treated with chloroformate **77** (124 mg, 0.4 mmol) in anhydrous CH₂Cl₂ (3 mL). After overnight stirring at room temperature, the solution was concentrated and the residue was coevaporated with anhydrous toluene (3 x 3 mL). The residue was diluted with water (15 mL) and the aqueous layer was back-extracted with CH₂Cl₂ (3x 50 mL). The organic layers were combined, dried and concentrated. The crude was purified via flash column chromatography (0-10% MeOH/DCM) to yield the title compound **83** (48 mg, 31%) as an orange solid; m.p. 105 °C to 108 °C. ¹H NMR (600 MHz, DMSO-d₆) δ 11.32 (br, 1 H, NH), 7.46 (d, 1 H, *J*=7.4 Hz, *H*-6), 7.45 (s, 1 H, *H*-6''), 6.69 (dd, 1 H, *J*=9.1, 2.6 Hz, *H*-9''), 6.55 (d, 1 H, *J*=2.5 Hz, *H*-7''), 6.20 (t, 1 H, *J*=6.9 Hz, *H*-1'), 5.97 (s, 1 H, *H*-3''), 5.47 (d, 1 H, *J*=4.4 Hz, 3'-OH), 5.37 (d, 1 H, *J*=4.4 Hz, *H*-13''), 4.39-4.33 (m, 2 H, *H*-5'), 4.26 (ddt, 1 H, *J*=6.5, 4.4, 3.8 Hz, *H*-4'), 3.95 (dt, 1 H, *J*=5.8, 3.8 Hz, *H*-4'), 3.43 (q, 4 H, *J*=7.1 Hz, *H*-11''), 2.18 (dt, 1 H, *J*=13.7, 6.9 Hz, *H*-2'a), 2.10 (ddd, 1 H, *J*=13.6, 6.6, 3.9 Hz, *H*-2'b), 1.71 (d, 1 H, *J*=1.1 Hz, *H*-7), 1.11 (t, 6 H, *J*=7.0 Hz, *H*-12''); ¹³C NMR (150 MHz, DMSO-d₆) δ 163.67 (C-4), 160.53 (C-2''), 155.88 (C-10''), 154.02 (C-14''), 150.54 (C-2), 150.41 (C-8''), 149.79 (C-4''), 135.85 (C-6), 125.60 (C-6''), 109.81 (C-5), 108.80 (C-7''), 105.42 (C-3''), 105.11 (C-5''), 96.88 (C-9''), 83.87 (C-1'), 83.37 (C-4'), 70.06 (C-3'), 67.91 (C-5'), 64.85 (C-13''), 44.04 (C-11''), 36.59 (C-2'), 12.34 (C-7), 12.11 (C-12''); IR *v*_{max}(solid): 3412 (O-H), 2961 (N-H), 1752 (C=O DEACMOC), 1681 (C=O Aza), 1257 (C-O), 1076 (C-O) cm⁻¹; HRMS (ESI) calcd for C₂₅H₃₀N₃O₉ [M+H]⁺ 516.1982, found 516.1986.

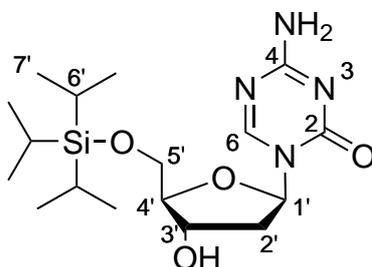
[(7-(diethylamino)-2-oxochroman-4-yl)methyl(perfluorophenyl)] carbonate (DEACMOC-PFP) **90**

A solution of 7-(diethylamino)-4-(hydroxymethyl)chroman-2-one **76** (300 mg, 1.2 mmol) in dry THF (10 mL) was treated with DIPEA (261 μ L, 1.5 mmol) and stirred for 5 min at 0°C. Bis(pentafluorophenyl)carbonate (956 mg, 2.4 mmol) was added and stirring was continued at room temperature under the exclusion of light overnight. The reaction mixture was transferred onto a mixture of ethyl acetate and water (1:1, 200 mL) and the layers were separated and the organic layer was dried over MgSO_4 and concentrated *in vacuo*. The crude was purified *via* flash column chromatography (neat CH_2Cl_2) to obtain DEACMOC-PFP **90** as a yellow oil (518 mg, 94%); m.p. 113°C to 114°C; ^1H NMR (600 MHz, CDCl_3) δ 7.29 (d, 1 H, $J=9.0$ Hz $H-6$), 6.61 (dd, 1 H, $J=9.0, 2.6$ Hz, $H-7$), 6.53 (d, 1 H, $J=2.7$ Hz, $H-9$), 6.22 (t, 1 H, $J=1.1$ Hz, $H-3$), 5.44 (d, 1 H, $J=1.1$ Hz, $H-13$), 3.43 (q, 4 H, $J=7.2$ Hz, $H-11$), 1.22 (t, 6 H, $J=7.1$ Hz, $H-12$); ^{13}C NMR (150 MHz, CDCl_3) δ 161.97 (C-2), 156.45 (C-10), 151.18 (C-14), 151.05 (C-8), 147.41 (C-4), 124.35 (C-6), 109.02 (C-7), 106.95 (C-3), 105.53 (C-5), 98.00 (C-9), 66.97 (C-13), 44.93 (C-11), 12.52 (C-12); IR v_{max} (solid): 1782 (C=O), 1722 (C=O), 1513 (C=C), 1234 (C-F) cm^{-1} LRMS (ESI) calcd for $\text{C}_{21}\text{H}_{16}\text{F}_5\text{NO}_5$ $[\text{M}+\text{H}]^+$ 457.3530, found 457.3511.

2'-Deoxy-5'-[7-(diethylamino)-2-oxochromn-4-yl]methoxycarbonyl]-5-azacytidine (5'-DEACMOC-dAC) **84**

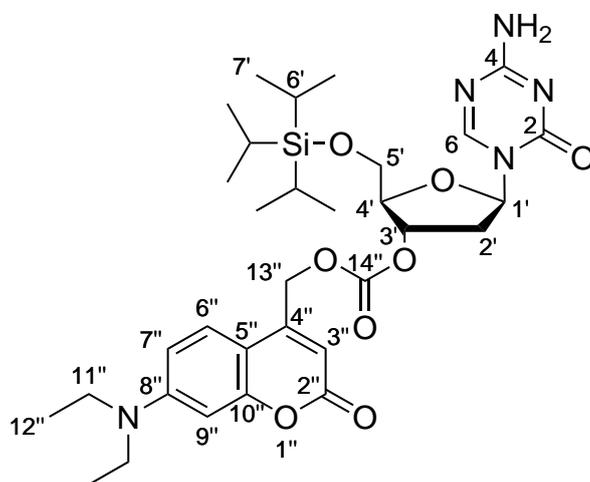


5-aza-2'-deoxycytidine **6** (157 mg, 0.69 mmol), DEACMOC-PFP **90** (518 mg, 1.13 mmol) and DMAP (4 mg, 0.03 mmol, 0.04 eq) were dissolved in anhydrous DMF (6 mL) and stirred at room temperature for 24 h under inert atmosphere. After 24 h, the reaction mixture was concentrated and coevaporated with anhydrous toluene (3 x 3 mL). The residue was dissolved in CH₂Cl₂ (50 mL) and washed with water. The organic layer was dried and concentrated *in vacuo*. The crude onto was dry-loaded onto and purified *via* flash column chromatography (0-10% MeOH/CH₂Cl₂) to yield the title compound **84** (76 mg, 22%) as a yellow solid. m.p. 145 °C to 147 °C ¹H NMR (600 MHz, DMSO-d₆) δ 8.30 (s, 1 H, *H*-6), 7.59 (d, 2 H, *J*=11.0 Hz, *aza*-NH₂), 7.47 (dd, 1 H, *J*=14.2, 9.1 Hz, *H*-6''), 6.71 (ddd, *J*=10.5, 9.1, 2.6 Hz, *H*-7''), 6.55 (t, 1 H, *J*=2.8 Hz, *H*-9''), 6.04 (t, 1 H, *J*=6.7 Hz, *H*-1'), 6.01 (d, 1 H, *J*=2.5 Hz, *H*-3''), 5.46 (d, 1 H, *J*=4.5 Hz, 3'-OH), 5.36 (d, 2 H, *J*=3.2 Hz, *H*-13''), 4.39 (dd, 1 H, *J*=11.6, 3.7 Hz, *H*-5'a), 4.34 (dd, 1 H, *J*=11.6, 6.4 Hz, *H*-5'b), 4.37 (dq, 1 H, *J*=6.6 Hz, *H*-3'), 4.00 (dt, 1 H, *J*=6.3, 4.1 Hz, *H*-4'), 3.43 (q, 4 H, *J*=7.1, 2.6 Hz, *H*-11''), 2.30 (dt, 1 H, *J*=13.6, 6.6 Hz, *H*-2'a), 2.21 (ddd, 1 H, *J*=13.6, 6.7, 4.6 Hz, *H*-2'b), 1.12 (td, 6 H, *J*=7.1, 2.5 Hz, *H*-12''); ¹³C NMR (150 MHz, DMSO-d₆) δ 165.85 (C-2), 160.66 (C-2''), 156.13 (C-10''), 155.86 (C-6), 153.97 (C-4), 153.44 (C-14''), 152.94 (C-4''), 150.56 (C-8''), 125.60 (C-6''), 108.85 (C-7''), 105.33 (C-5''), 105.10 (C-3''), 96.88 (C-9''), 84.94 (C-1'), 83.90 (C-4'), 70.06 (C-3'), 67.93 (C-5'), 64.83 (C-13''), 44.04 (C-11''), 40.07 (C-2'), 12.34 (C-12''); IR *v*_{max}(solid): 3392 (O-H), 3211 (N-H), 3085 (N-H), 1758 (C=O DEACMOC), 1711 (C=O Ar DEACMOC), 1682 (C=O Ar), 1599 (C=C), 1255 (C-O Ar), 1197 (C-O), 1048 (C-O) cm⁻¹; HRMS (ESI) calcd for C₂₃H₂₈N₅O₈ [M+H]⁺ 502.1938, found 502.1917.

2'-Deoxy-5'-(triisopropylsilyl)-5-azacytidine (5'-TIPS-dAC) 91

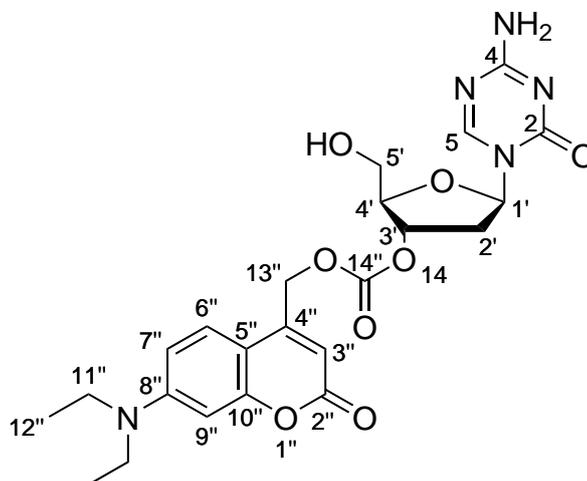
To a stirred, cooled solution of 5-aza-2'-deoxycytidine **6** (150 mg, 0.66 mmol), imidazole (98 mg, 1.45 mmol) and DMAP (8 mg, 0.07 mmol) in anhydrous DMF (4 mL) was added a solution of triisopropyl silyl chloride (154 μ L, 0.72 mmol) in anhydrous DMF (2 mL). The mixture was stirred for 1 h at 0 °C and overnight at room temperature. The crude was concentrated and purified via flash column chromatography (0-5% MeOH/CH₂Cl₂) to obtain the title compound **91** as a white solid (133 mg, 52%); m.p. 179 °C to 180 °C; ¹H NMR (600 MHz, DMSO-d₆) δ 8.38 (s, 1 H, *H*-6), 7.51 (d, 2 H, *J*=14.0 Hz, -NH₂), 6.04 (t, 1 H, *J*=6.4 Hz, *H*-1'), 5.28 (d, 1 H, *J*=4.4 Hz, 3'-OH), 4.27 (m, 1 H, *H*-3'), 3.91-3.87 (m, 1 H, *H*-5'a), 3.88-3.86 (m, 1 H, *H*-4'), 3.83-3.80 (m, 1 H, *H*-5'b), 2.26 (ddd, 1 H, *J*=13.5, 6.6, 4.1 Hz, *H*-2'a), 2.21 (ddd, 1 H, *J*=13.6, 6.4, 6.4 Hz, *H*-2'b), 1.14-1.08 (m, 3 H, *H*-6'), 1.04 (d, 18 H, *J*=7.0 Hz, *H*-7'); ¹³C NMR (150 MHz, DMSO-d₆) δ 165.96 (C-2), 155.38 (C-6), 153.03 (C-4), 84.30 (C-1'), 85.14 (C-4'), 69.78 (C-3'), 63.10 (C-5'), 40.63 (C-2'), 17.67 (C-7'), 11.34 (C-6'); IR ν_{max} (solid): 3343 (O-H), 2942 (N-H), 2865 (N-H), 1638 (C=O Aza), 1289 (Si-C), 1090 (Si-O), 1062 (C-O), 1008 (C-O) cm⁻¹; HRMS (ESI) calcd for C₁₇H₃₂N₄O₄Si [M+H]⁺ 384.21928, found 384.55200.

2'-Deoxy-3'-(7-(diethylamino)-2-oxochromn-4-yl)methoxycarbonyl)-5'-(triisopropylsilyl)-5'-azacytidine (3'-DEACMOC-5'-TIPS-dAC) **92**

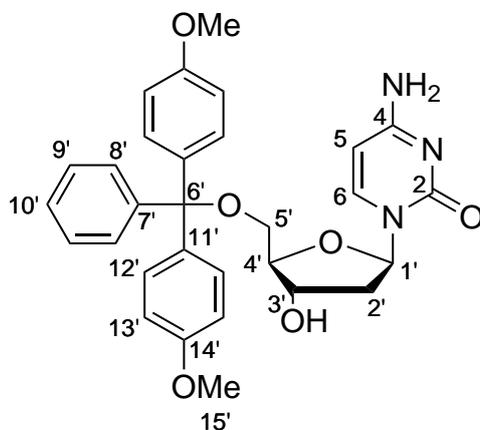


To a stirred solution of compound **91** (364 mg, 0.94 mmol) and DMAP (11 mg, 0.09 mmol) in anhydrous DMF (4 mL) was added a solution of **7** (340 mg, 0.74 mmol) in anhydrous DMF (4 mL) at 0 °C. The mixture was stirred for 6 h and was concentrated *in vacuo*. Traces of DMF were removed via coevaporation with toluene (3 x 3 mL). The crude was purified via flash column chromatography (0-5% MeOH/CH₂Cl₂) to obtain compound **92** as a yellow solid (191 mg, 31%); m.p. 82 °C to 84 °C; ¹H NMR (600 MHz, DMSO-d₆) δ 8.35 (s, 1 H, *H*-6), 7.62 (d, 2 H, *J*=20.1 Hz, aza-NH₂), 7.48 (d, 1 H, *J*=9.1 Hz, *H*-6''), 6.71 (dd, *J*=9.2, 2.6 Hz, *H*-7''), 6.56 (d, 1 H, *J*=2.5 Hz, *H*-9''), 6.05 (dd, 1 H, *J*=7.8, 6.2 Hz, *H*-1'), 6.02 (s, 1 H, *H*-3''), 5.36 (s, 2 H, *H*-13''), 5.24 (dt, 1 H, *J*=6.5, 2.3 Hz, *H*-3'), 4.25 (q, 1 H, *J*=3.0 Hz, *H*-4'), 3.92 (dd, 1 H, *J*=11.1, 3.75 Hz, *H*-5'a), 3.87 (dd, 1 H, *J*=11.2, 3.9 Hz, *H*-5'b), 3.43 (dt, 4 H, *J*=7.6, 6.8 Hz, *H*-11''), 2.58 (ddd, 1 H, *J*=14.3, 5.9, 2.4 Hz, *H*-2'a), 2.40 (ddd, 1 H, *J*=14.5, 7.6, 6.5 Hz, *H*-2'b), 1.12 (t, 6 H, *J*=7.0 Hz, *H*-12''), 1.12-1.05 (m, 3 H, *H*-6'), 1.01 (d, 18 H, *J*=6.9 Hz, *H*-7'); ¹³C NMR (150 MHz, DMSO-d₆) δ 165.90 (C-2), 160.56 (C-2''), 155.93 (C-10''), 155.43 (C-6), 153.36 (C-4), 153.36 (C-14''), 152.86 (C-4''), 150.55 (C-8''), 125.72 (C-6''), 108.82 (C-7''), 105.84 (C-5''), 105.21 (C-3''), 96.87 (C-9''), 85.48 (C-1'), 84.60 (C-4'), 65.06 (C-3'), 63.83 (C-5'), 63.39 (C-13''), 44.04 (C-11''), 39.38 (C-2'), 17.61 (C-6'), 12.34 (C-12''), 11.26 (C-7'); IR *v*_{max}(solid): 2941 (N-H), 2865 (N-H), 1711 (C=O), 1600 (C=O), 1246 (Si-C), 1196 (C-O), 1126 (C-O), 1065 (Si-O) cm⁻¹; HRMS (ESI) calcd for C₃₂H₄₇N₅O₈Si [M+H]⁺ 658.3267, found 658.3268.

2'-Deoxy-3'-(7-(diethylamino)-2-oxochromn-4-yl)methoxycarbonyl)-5-azacytidine (3'-DEACMOC-dAC) **93**



To a stirred solution of compound **92** (87 mg, 0.13 mmol) in dry methanol (3 mL), acetyl chloride (20 μ L, 0.28 mmol) was added under inert atmosphere. The mixture was stirred at room temperature for 3 h in the dark. It was then partially concentrated and diluted with CH₂Cl₂ (20 mL) and neutralised with 10% sodium bicarbonate solution. The organic layer was washed with water (10 mL), dried and concentrated. The crude was purified via flash column chromatography (0-8% MeOH/CH₂Cl₂) to obtain compound **93** (36.4 mg, 56%) as a yellow solid; m.p. 167 °C to 170 °C ¹H NMR (600 MHz, DMSO-d₆) δ 8.47 (s, 1 H, *H*-6), 7.59 (dd, 2 H, *J*=13.7, 2.3 Hz, aza-NH₂), 6.72 (dd, 1 H, *J*=9.1, 2.6 Hz, *H*-6''), 6.56 (dd, *J*=9.2, 2.6 Hz, *H*-7''), 6.56 (d, 1 H, *J*=2.5 Hz, *H*-9''), 6.05 (dd, 1 H, *J*=8.1, 6.0 Hz, *H*-1'), 6.02 (s, 1 H, *H*-3''), 5.36 (s, 2 H, *H*-13''), 5.22 (t, 1 H, *J*=5.3 Hz, *H*-3'), 4.17 (q, 1 H, *J*=3.8, 2.0 Hz, *H*-4'), 3.64 (m, 2 H, *H*-5'), 3.43 (q, 4 H, *J*=7.1 Hz, *H*-11''), 2.43 (ddd, 1 H, *J*=14.3, 8.3, 6.1 Hz, *H*-2'b), 1.12 (t, 6 H, *J*=7.1 Hz, *H*-12''); ¹³C NMR (150 MHz, DMSO-d₆) δ 165.66 (C-2), 160.59 (C-2''), 156.02 (C-10''), 155.90 (C-6), 153.44 (C-4), 153.03 (C-14''), 150.56 (C-4''), 149.68 (C-8''), 125.66 (C-6''), 106.66 (C-7''), 105.66 (C-5''), 105.18 (C-3''), 96.88 (C-9''), 85.40 (C-1'), 84.94 (C-4'), 79.26 (C-3'), 64.95 (C-5'), 61.25 (C-13''), 44.05 (C-11''), 37.53 (C-2''), 12.34 (C-12''); IR *v*_{max}(solid): 3433 (O-H), 3081 (N-H), 2969 (N-H), 1711 (C=O), 1666 (C=O), 1600 (C=C), 1252 (C-N), 1196 (C-O), 1098 (C-O) cm⁻¹; HRMS (ESI) calcd for C₂₃H₂₇N₅O₈ [M+H]⁺ 502.1932, found 502.1929.

2'-Deoxy-5'-dimethoxytrityl-azacytidine²⁵⁵ **97**

To a stirred solution of dAC **6** (100 mg, 0.44 mmol) in dry pyridine (3 mL), dimethoxytrityl chloride (153 mg, 0.45 mmol) was added under inert atmosphere. The mixture was stirred at room temperature overnight in the dark. The solution was then concentrated and the crude was purified via flash chromatography (0-7% MeOH/CH₂Cl₂) to obtain compound **97** (122 mg, 52%) as a white solid. ¹H NMR (600 MHz, DMSO-d₆) δ 8.56 (s, 1 H, *H*-6), 7.39-7.21 (m, 9H, *H*-8', *H*-9', *H*-10'), 6.88 (d, 2 H, *J*=8.9 Hz, -NH₂), 6.04 (t, 1 H, *J*=6.4 Hz, *H*-1'), 5.33 (d, 1H, *J*=4.4 Hz, *H*-5), 4.25-4.22 (m, 1H, *H*-4'), 3.94-3.92 (m, 1H, *H*-3'), 3.73 (s, 6H, *H*-15'), 3.22-3.14 (m, 2H, *H*-5'), 2.25-2.23 (m, 2H, *H*-2'); ¹³C NMR (150 MHz, DMSO-d₆) δ 165.52 (C-4), 157.36 (C-14'), 155.24 (C-2), 143.64 (C-6), 143.53 (C-7'), 136.25 (C-11'), 129.2-128.3 (C-8', C-9', C-10', C-12'), 113.53 (C-13'), 94.32 (C-5), 93.25 (C-6'), 87.94 (C-4'), 70.34 (C-3'), 63.24 (C-5'), 54.25 (C-15'), 40.43 (C-2'); Measurements are in line with literature reported values.²⁵⁵

A.3 Cell biology experiments

A.3.1 Cell lines and reagents

The Saos-2 human osteosarcoma cell line was obtained from the laboratory of Dr Sandra Strauss (UCL Cancer Institute). The T24 human bladder, A549 human lung cell carcinoma and HEK293T human embryonic kidney cell lines were received from Dr Andrew Feber and Dr Anna Köferle (UCL Cancer Institute). All cell lines were grown as adherent monolayers in respected cell culture media, supplemented with 10% foetal bovine serum (Gibco, #10500-064) on uncoated tissue culture plasticware. Saos-2, T24 and A549 cells were grown in RPMI-1640 medium (Lonza, #12-702F) and HEK293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM, Gibco, #12491-015). All cells were maintained at 37 °C with 5% CO₂.

A.3.2 Cell culture

Cell passaging

30 min prior to passaging, aliquots of Trypsin-EDTA (Sigma, #T4049) and the relevant growth media were prewarmed in a 37 °C water bath and Dubbleco's Phosphate-buffered Saline (DBPS, Gibco, #14190136) solution was allowed to reach room temperature. A small volume of old media was aliquoted into a 15 mL falcon tube (2 mL from T25 flasks, 7 mL from T75 and 10 mL from T125 flasks). After aspiration of the leftover growth media, the plate containing the cells was gently washed with DBPS, avoiding to disrupt the cell layer. The remaining DBPS was aspirated and pre-warmed trypsin (1 mL for T25 flask, 4 mL for T75 flask and 10 mL for T125 flask) was added. The trypsin was evenly distributed by gently tilting the flask and the flask was placed into an incubator (37 °C, 5% CO₂) for a maximum of 5 min.

The detached cells were then resuspended in the old media (that was aliquoted previously) and the suspension was transferred to a new 15 mL falcon tube (50 mL in case the volume exceeded 12 mL). The tubes were spun down for 3 min at 1200 rpm to pellet the cells. The remaining supernatant was discarded and the pellet was resuspended in fresh growth media. The suspension of cells was then split 1:5 (depending on their confluency and growth rate) in new flasks (T25, T75 or T125) and

pre-warmed growth media were added for the appropriate flasks size (5 mL for T25, 15 mL for T75 and 25 mL for T125 flasks).

Cell freezing

Cells were detached and pelleted as described in the previous section. Instead of adding fresh growth medium, the pellets were resuspended in freezing medium (10% DMSO to growth medium). 0.8 mL of the resulting cell suspension was transferred to a cryovial and was immediately placed on dry ice for 10 min before being transferred to the -80°C freezer overnight and to liquid nitrogen for long-term storage.

Cell counting

In order to calculate accurate seeding densities, cell numbers were determined using a Vi-Cell XR (Beckman Coulter) automated cell viability analyser. This method uses the trypan blue exclusion method to count cells within a $500\ \mu\text{L}$ volume.

A.3.3 Photocaged azanucleoside treatments

Preparation of treatment stocks

Photocaged azanucleosides (3'-DEACMOC dAC **93** or 5'-DEACMOC-dAC **84**) were dried under high vacuum for 2 x 8 h to remove any residual solvent. Stocks of photocaged analogues and commercially available dAC were prepared at 10 mM in DMSO (between 500 and 1000 μL , in MS vials). The concentration of each stock was confirmed by HPLC. The stocks were stored at 4°C until further dilution for cell studies.

Preparation of cells, treatment and irradiation

Cells were seeded into 10 cm TC-treated plates (Corning, #430167) at an initial density of 1×10^5 cells/plate. For each treatment set per cell line, at least 36 cell plates were prepared to account for the different concentrations (0.1 - 1.5 μM), analogues (dAC, photocaged dAC), conditions (irradiation or control) and triplicates (Table A.2).

The cells were allowed to grow to about 70-80% confluency before the treatment administration 72 h after seeding.

Table A.2 outlines the overall treatment plan for one set of cell experiments for one cell line. Groups A/B/C corresponded to triplicates. Due to the limited space in the UV irradiating cabinet (4 plates), groups A/B/C were treated and irradiated 1 h apart.

Plate no.	Group A	Group B	Group C
1	Media UV	Media UV	Media UV
2	0.1 μ M CD UV	0.1 μ M CD UV	0.1 μ M CD UV
3	0.5 μ M CD UV	0.5 μ M CD UV	0.5 μ M CD UV
4	1.5 μ M CD UV	1.5 μ M CD UV	1.5 μ M CD UV
5	Media/DMSO	Media/DMSO	Media/DMSO
6	0.1 μ M CD	0.1 μ M CD	0.1 μ M CD
7	0.5 μ M CD	0.5 μ M CD	0.5 μ M CD
8	1.5 μ M CD	1.5 μ M CD	1.5 μ M CD
9	Media	Media	Media
10	0.1 μ M dAC	0.1 μ M dAC	0.1 μ M dAC
11	0.5 μ M dAC	0.5 μ M dAC	0.5 μ M dAC
12	1.5 μ M dAC	1.5 μ M dAC	1.5 μ M dAC

Table A.2 Plate setup for dAC and photocaged dAC cell treatments

The following process was repeated three times for each treatment group: 30 min prior to treatment, the relevant cell culture media (RPMI-1640 or DMEM) was warmed to 37°C in a water bath. Treatment dilutions were prepared in 50 mL falcon tubes as described in Table A.3.

Tube no.	Treatment condition	Stock/solution added	Fresh Media (mL)	Final Volume (mL)
1	DMSO/Media	4.5 μ L DMSO	30	30
2	1.5 μ M dAC	4.5 μ L of 10 mM dAC stock	30	21
3	0.5 μ M dAC	9 mL of 1.5 μ M dAC (2)	18	22
4	0.1 μ M dAC	5 mL of 0.5 μ M dAC (3)	20	25
5	1.5 μ M CD	4.5 μ L of 10 mM CD stock	30	21
6	0.5 μ M CD	9 mL of 1.5 μ M CD (5)	18	22
7	0.1 μ M CD	5 mL of 0.5 μ M CD (6)	20	25

Table A.3 Preparation of treatment dilutions for cell experiments

After the dilutions were prepared, the plates were taken out of the 37 °C incubator and the old media was aspirated away. The prepared treatment solutions were placed into each plate (10 mL per plate) and plates 1-4 (refer to table A.3) were placed in the UV cabinet and irradiated for 1 h at $\lambda = 365$ nm. The remaining plates (5-12) were placed back into the incubator. After 1 h, the irradiated plates (1-4) were also placed back into the incubator.

The cells were harvested 24 h after treatment. Aliquots of the old media (7 mL) were withdrawn from each plate. The remaining old media was aspirated away and the cells were washed with DPBS (7 mL). After removal of the DPBS, pre-warmed trypsin was added (3 mL) and the plates were incubated for 1-3 min at 37 °C. The trypsinised cells were transferred to 15 mL falcon tubes together with the 7 mL old media aliquots taken previously and were spun down for 4 min at 1200 rpm. The supernatant was discarded and the cell pellets were used immediately or were stored at -20 °C until further analysis.

Preparation of samples for global DNA Methylation analysis

Cell pellets from photocaged azanucleoside treatments were either used immediately or were defrosted to room temperature from -20 °C. Genomic DNA from cell pellets were purified using DNeasy Blood & Tissue Kit (Qiagen; #69504), according to manufacturer's instructions.

The cells were resuspended in PBS (200 μ L) and proteinase K (20 μ L) and RNase A (4 μ L, 100 mg/mL, Qiagen, #19101) were added and the samples were vortexed and incubated for 2 min at room temperature. Buffer AL (200 μ L) was added and the samples were thoroughly mixed incubated at 56 °C for 10 min with continuous rotation. Absolute ethanol (200 μ L) was added to each sample and after vortexing, the solutions were transferred into a DNeasy Mini spin column equipped with a collection tube. The samples were centrifuged at 8000 rpm for 1 min and the collection tube was discarded. The samples were subsequently washed with Buffer AW1 (500 μ L) and Buffer AW2 (500 μ L). After each buffer wash, the samples were centrifuged either at 8000 rpm for 1 min or 16000 rpm for 3 min. In order to maximise the amount of DNA extracted, we modified the centrifugation times to 3 min and 5 min respectively. As a final step, genomic DNA was eluted in RNase free water (100 μ L) and were collected in clean, labelled eppendorf tubes.

The DNA concentration was determined using the Qubit dsDNA HS Assay Kit following the manufacturer's protocol. The samples were then sent to Cambridge for LC-MS/MS analysis to determine the %5mC content of purified genomic DNA samples.

Preparation of samples for Western blot analysis

Cells were washed with cold PBS, counted, resuspended and lysed in an appropriate volume of lysis buffer (CellLytic M, #C2978 containing 50x Pierce Protease inhibitor, #88666) with continuous rotation for 15 min at 4°C. Samples were spun down for 15 min at 16 000 g, 4°C. The supernatant was transferred to a fresh Eppendorf tube and the protein concentration was determined by the Bradford Assay.

A.3.4 Global methylation analysis *via* LC-MS/MS

LC-MS/MS analysis was performed by Dr Shiqing Mao (Department of Chemistry, University of Cambridge) as follows: 500 ng of genomic DNA was incubated with 5 U of DNA Degradase Plus (Zymo Research; # E2020) at 37°C for 3 h. The resulting mixture of 2'-deoxynucleosides was analysed on a Triple Quad 6500 mass spectrometer (AB Sciex) fitted with an Infinity 1290 LC system (Agilent) and an Acquity UPLC HSS T3 column (Waters), using a gradient of water and acetonitrile with 0.1% formic acid. External calibration was performed using synthetic standards, and for accurate quantification, all samples and standards were spiked with isotopically labelled nucleosides.

A.3.5 Statistical analysis of methylation values

Cellular DNA methylation values (% 5mC) were expressed as mean \pm SD. To verify a between treatment difference, a two-tailed Student's t-test was performed.

A.3.6 Methylation analysis *via* the Infinium 450K BeadChip

For DNA methylation analysis, 500 ng of each genomic DNA (gDNA) sample was bisulfite converted using an EZ DNA Methylation kit (Zymo Research, #D5001) and eluted in 12 μ L of M-Elution buffer. 4 μ L were used as input per array of the Infinium 450K BeadChip and processed by the UCL Genomics microarray facility according to

the manufacturer's protocol. Raw data files were uploaded to R and analysed using the Chip Analysis Methylation Pipeline (ChAMP), available from Bioconductor. Probes were filtered for a detection p-value < 0.01, leaving 483,824 probes in the final dataset. Finally, the data was normalized using the SWAN method.

A.3.7 Reagents for western blot analysis

SDS PAGE Gel (8%) Running gel: 4.6 mL double-distilled (dd) H₂O, 2.7 mL 30% Bis-acrylamide solution, 37.5:1 (BioRad, #EC-890), 2.5 mL 1.5M Tris pH 8.8, 100 μ L 10% SDS solution, 100 μ L mL 10% Ammonium persulphate (APS) solution, 6 μ L TEMED (Scientific Laboratory, # EC-503)

Stacking gel: 2.1 mL ddH₂O, 0.5 mL Bis-acrylamide, 0.38 mL 0.5 M Tris-HCl pH 6.8, 30 μ L 10% SDS solution, 3 μ L 10% APS solution

5x Laemmli buffer (100 mL): 12.5 mL 1M Tris-HCl pH 6.8, 25 mL 20% SDS solution, 57.5 mL ddH₂O, 5 mL β -mercaptoethanol, bromophenol blue

10 x SDS Running buffer (2 L): 288.2 g Glycine (1.92 M), 60.6 g Tris-base (250 mM), 100 mL of 20% SDS solution, fill up to 2 L with ddH₂O

10 x Western blot transfer buffer (2 L): 288.2 g Glycine (1.92 M), 60.6 g Tris-base (250 mM), 100 mL of 20% SDS solution, add 20%(v/v) methanol (400 mL), fill up to 2 L with ddH₂O

10 x TBS pH 7.4 (2 L): 175.32 g NaCl (3 M), 121.14 g Tris-base (1 M), 75 mL of 37% HCl, fill up to 2 L with ddH₂O and adjust pH to 7.4

A.3.8 Western blot analysis

Samples were mixed with 5x Laemmli Buffer and incubated at 95 °C for 5 min. 30 μ g whole cell lysate of each sample was loaded on a 8% SDS PAGE gel. As a standard, 6 μ L of Precision Plus Protein Standard, 10-250 kD (BioRad, #1610373) was included. The gels were run in 1x SDS-PAGE running buffer at 35V for 30 min then 100V for 2 h.

Separated proteins were transferred from an SDS PAGE gel to a 0.45 μ m pore polyvinylidene (PVDF) membrane *via* wet transfer. PVDF membrane (Sigma, #Z671010) was cut to the correct size of the gel and was prepared for the transfer by soaking in

methanol for 30 seconds. Sponges and blotting paper were soaked in cold transfer buffer. Gels were removed from the electrophoresis tank and removed from the glass plates. They were sandwiched in a transfer cassette as shown in FIG A.2. The transfer cassettes were loaded into the gel tank with an ice pack and cold transfer buffer supplemented with 0.1% SDS (0.5 mL). Electrophoresis was performed at 60V for 2h at 4 °C.

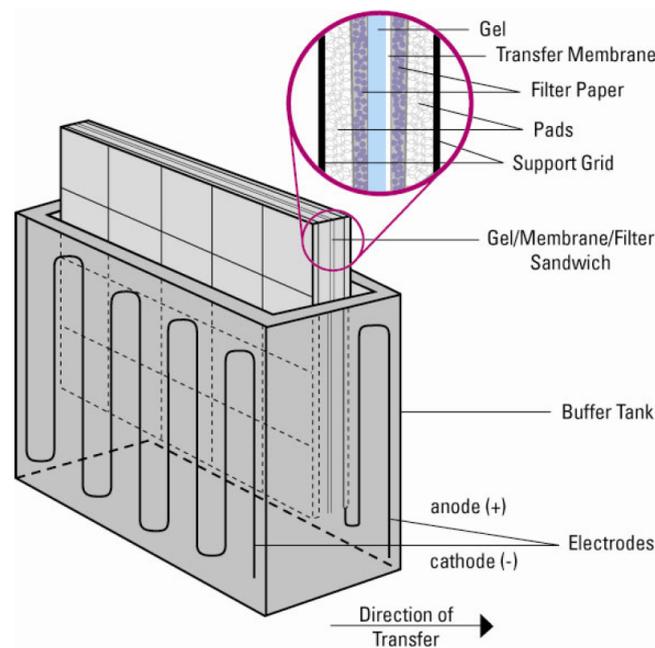


Fig. A.2 Electrophoretic transfer setup

The membrane was blocked with 5% Milk/TBST for 1 h at 4 °C. Membranes were incubated with primary antibodies (1:200 anti-DNMT1, #sc-271729 and 1:300 anti-GAPDH, #MAB374) overnight at 4 °C on a rocker, followed by 3 x 10 min rinse in TBST. Membranes were then incubated with the relevant secondary antibodies (1:4000, HRP-linked Mouse IgG, #NAV931V) for 1 h at room temperature on a rocker, followed by a 5 x 10 min TBST wash.

Visualisation was achieved by the addition of 1:1 luminol and hydrogen peroxide mixture (2 mL) (SuperSignalTM West Femto, 34095, Thermo). The blot was incubated in the dark for 10 min. Chemoluminescence was detected using the ODYSSEY CLx Imaging system (LI-COR Biosciences, UK)

Densitometric analysis of western blot bands

In order to semi-quantify the amount of DNMT1 in the western blot analyses, the blots were subjected to densitometric quantification. The data are expressed as arbitrary units obtained by analysing the bands by using the software Image J. The intensities of the DNMT1 and GAPDH bands were measured and the DNMT1 bands were normalised to GAPDH bands. Tables A.4 - A.7 summarise the normalised DNMT1 band intensities and calculated relative baseline intensities in 3'- and 5'-CD treated Saos-2 and T24 cells. Figures A.3 - A.6 visually summarise the change in relative band intensities against different treatment conditions.

Lane no.	Treatment condition	Normalised DNMT1	% Baseline intensity
1	Media (M)	9755	100
2	M/DMSO	9666	99
3	0.1 μ M dAC	1696	17
4	0.5 μ M dAC	516	5
5	1.5 μ M dAC	1391	14
6	M-UV	12001	100
7	M/DMSO-UV	11814	98
8	0.1 μ M 3'-CD UV	2865	24
9	0.5 μ M 3'-CD UV	419	3
10	1.5 μ M 3'-CD UV	1449	12
11	M/DMSO	6232	100
12	0.1 μ M 3'-CD	1576	25
13	0.5 μ M 3'-CD	665	11
14	1.5 μ M 3'-CD	445	7

Table A.4 Normalised DNMT1 band intensities in 3'-CD treated Saos-2 cells

Lane no.	Treatment condition	Normalised DNMT1	% Baseline intensity
1	M/DMSO	9856	100
2	0.1 μ M dAC	3383	34
3	0.5 μ M dAC	2834	29
4	1.5 μ M dAC	3772	38
5	M-UV	13052	100
6	M/DMSO-UV	12613	97
7	0.1 μ M 5'-CD UV	5151	39
8	0.5 μ M 5'-CD UV	3438	26
9	1.5 μ M 5'-CD UV	3117	24
10	M/DMSO	14113	100
11	0.1 μ M 5'-CD	5702	40
12	0.5 μ M 5'-CD	3927	28
13	1.5 μ M 5'-CD	2327	16

Table A.5 Normalised DNMT1 band intensities in 5'-CD treated Saos-2 cells

Lane no.	Treatment condition	Normalised DNMT1	% Baseline intensity
1	Media (M)	14569	100
2	M/DMSO	12102	83
3	0.1 μ M dAC	1245	9
4	0.5 μ M dAC	787	5
5	1.5 μ M dAC	804	6
6	M-UV	12237	100
7	M/DMSO-UV	12120	99
8	0.1 μ M 3'-CD UV	2251	18
9	0.5 μ M 3'-CD UV	718	6
10	1.5 μ M 3'-CD UV	1046	9
11	M/DMSO	9495	100
12	0.1 μ M 3'-CD	6374	67
13	0.5 μ M 3'-CD	1815	19
14	1.5 μ M 3'-CD	789	8

Table A.6 Normalised DNMT1 band intensities in 3'-CD treated T24 cells

Lane no.	Treatment condition	Normalised DNMT1	% Baseline intensity
1	M/DMSO	13326	100
2	0.1 μ M dAC	1773	13
3	0.5 μ M dAC	1045	8
4	1.5 μ M dAC	1282	10
5	M-UV	13917	100
6	M/DMSO-UV	15564	112
7	0.1 μ M 5'-CD UV	4118	30
8	0.5 μ M 5'-CD UV	1696	12
9	1.5 μ M 5'-CD UV	1898	14
10	M/DMSO	11706	100
11	0.1 μ M 5'-CD	5259	45
12	0.5 μ M 5'-CD	1674	14
13	1.5 μ M 5'-CD	1112	10

Table A.7 Normalised DNMT1 band intensities in 5'-CD treated T24 cells

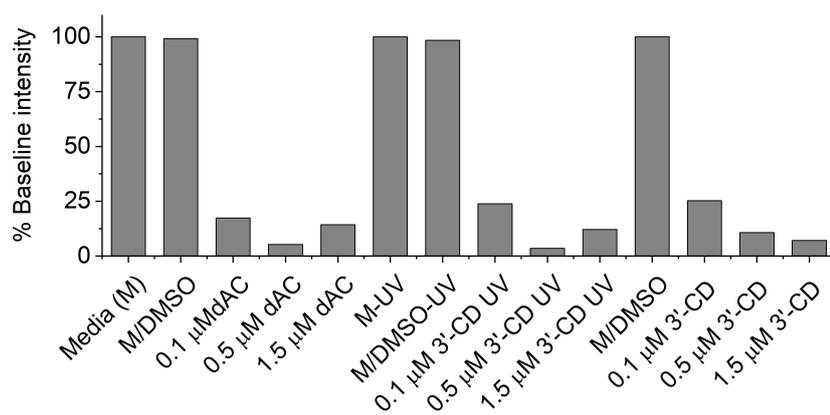


Fig. A.3 Relative DNMT1 band intensities in 3'-CD treated Saos-2 cells

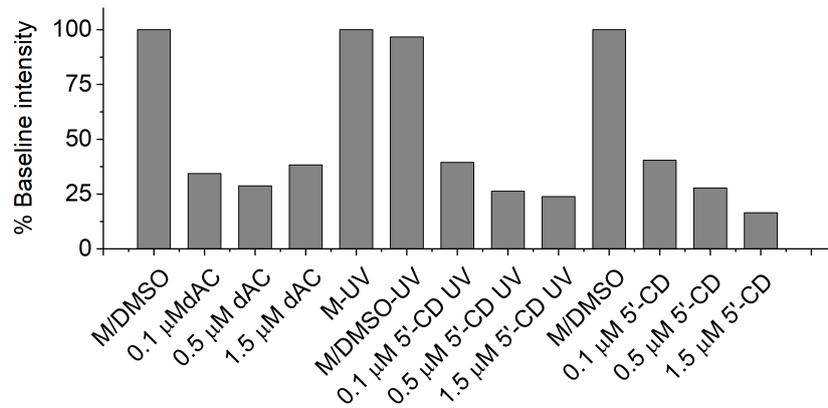


Fig. A.4 Relative DNMT1 band intensities in 5'-CD treated Saos-2 cells

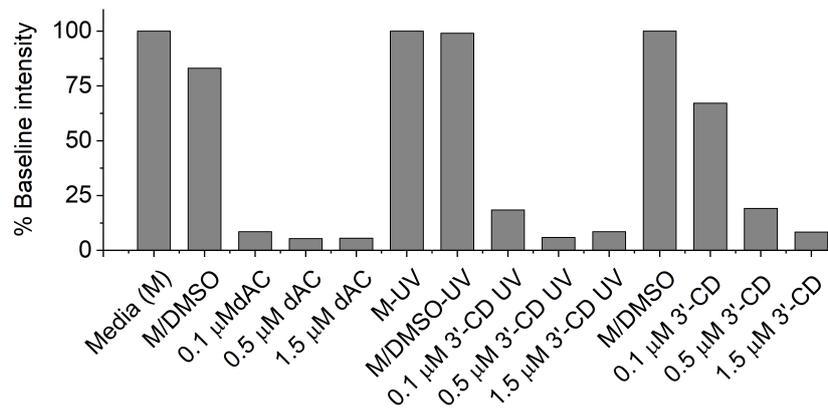


Fig. A.5 Relative DNMT1 band intensities in 3'-CD treated T24 cells

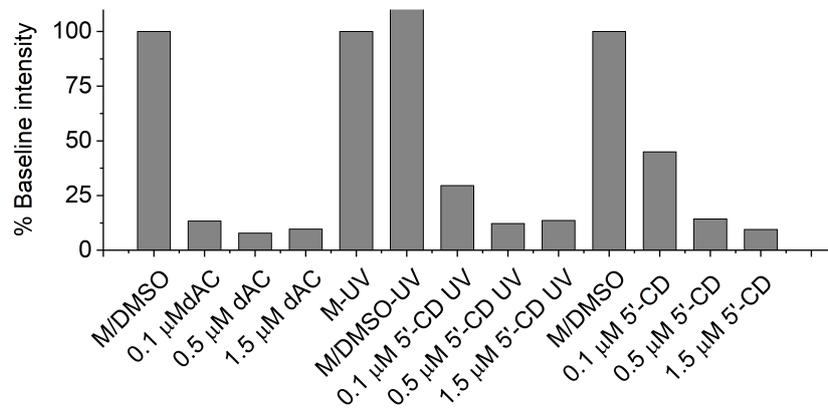


Fig. A.6 Relative DNMT1 band intensities in 5'-CD treated T24 cells

A.3.9 Cell viability assays

Low concentration treatments

Cells were seeded in 24-well tissue culture plates (BD Falcon #353047 or Corning Primaria #10288830) as described in Table A.8. For each cell type, two plates were prepared, one for illumination and one for dark control. Accurate cell counts were determined using a Vi-CELL™ XR cell viability analyzer (Beckman Coulter).

Cell type	Seeding density (cells/well)	Culture plate type
Saos-2	9.3×10^4	BD Falcon
T24	1.0×10^4	BD Falcon
HEK293T	2.5×10^4	Corning Primaria

Table A.8 Cell seeding densities for low concentration treatments

After seeding, the plates were transferred to an IncuCyte™ (Essen Bioscience) for automated monitoring of cell growth, allowing confluence measurements to be made at 1 or 2-h intervals over the entire course of the assay. 72 h after seeding, the cells were treated with dAC and 3'-caged dAC (3'-CD). dAC and photocaged dAC treatment dilutions were prepared in 15 mL falcon tubes 30 min prior to treatment (Table A.9).

Tube no.	Treatment solution	Stock/solution added	Fresh Media (mL)	Final Volume (mL)
1	DMSO/Media	1.8 μ L DMSO	12	12
2	1.5 μ M dAC	1.8 μ L of 10 mM dAC stock	12	8
3	0.5 μ M dAC	4 mL of 1.5 μ M dAC (2)	8	6
4	0.1 μ M dAC	2 mL of 0.5 μ M dAC (3)	8	10
5	1.5 μ M CD	1.8 μ L of 10 mM CD stock	12	8
6	0.5 μ M CD	4 mL of 1.5 μ M CD (5)	8	6
7	0.1 μ M CD	2 mL of 0.5 μ M CD (6)	8	10

Table A.9 Preparation of dAC and photocaged dAC dilutions for low concentration treatments

The old media was aspirated away and 500 μ L of the appropriate treatment solutions were added to each well. The treatment setup for the 24-well plates is shown in Table A.10. Each solution was administered in triplicates. One plate was placed in the UV

cabinet and was irradiated at $\lambda = 365$ nm for 1 h. The other plate was kept in the cell culture hood at room temperature and was covered with aluminium foil. After the completion of the irradiation, the plates were returned to the IncuCyte. 24 h after treatment, the plates were removed from the incucyte and the cell viability data was analysed.

	1	2	3	4	5	6
A	0.1 μ M 3'-CD (3x)			0.1 μ M dAC (3x)		
B	0.5 μ M 3'-CD (3x)			0.5 μ M dAC (3x)		
C	1.5 μ M 3'-CD (3x)			1.5 μ M dAC (3x)		
D	1.5 μ M DMSO (3x)			Media (3x)		

Table A.10 Cell viability treatment setup I.

High concentration treatments

Saos-2 cells were seeded in two 24-well tissue culture plates at a density of 2×10^4 cells/well. The plates were transferred to the IncuCyte for automated monitoring of cell growth. Three treatments were administered at 24-h intervals, starting 24 h after seeding of the cells. 30 min prior to each treatment, dilutions of 3'-CD and dAC at 1.5, 4.5 and 13.5 μ M were prepared in 15 mL falcon tubes using pre-warmed fresh media (Table A.11).

Tube no.	Treatment solution	Stock/solution added	Fresh Media (mL)	Final Volume (mL)
1	162 μ M dAC stock	48.6 μ L of 10 mM dAC stock	3	3
2	13.5 μ M dAC	1 mL of 162 μ M dAC stock (1)	11	8
3	4.5 μ M dAC	4 mL of 13.5 μ M dAC (2)	8	8
4	1.5 μ M dAC	4 mL of 4.5 μ M dAC (3)	8	12
5	162 μ M CD stock	48.6 μ L of 10 mM CD stock	3	3
6	13.5 μ M CD	1 mL of 162 μ M CD stock (5)	11	8
7	4.5 μ M CD	4 mL of 13.5 μ M CD (6)	8	8
8	1.5 μ M CD	4 mL of 4.5 μ M CD (7)	8	12

Table A.11 Preparation of dAC and photocaged dAC dilutions for for high concentration treatments

The plates were removed from the IncuCyte and the old media was aspirated away. 500 μL of treatment solution was added to each well following the setup in Table A.12. One plate was placed into the UV cabinet for 1 h irradiation at $\lambda = 365$ nm and the other plate was covered in aluminium foil and was kept in the tissue culture hood at room temperature. After the irradiation was complete, the plates were returned to the IncuCyte for imaging. The treatment process was carried out for three times in total.

	1	2	3	4	5	6
A	1.5 μM 3'-CD (3x)			1.5 μM dAC (3x)		
B	4.5 μM 3'-CD (3x)			4.5 μM dAC (3x)		
C	13.5 μM 3'-CD (3x)			13.5 μM dAC (3x)		
D	Media (6x)					

Table A.12 Cell viability treatment setup II.

A.3.10 ELISA assays

ELISA assays were performed according to the manufacturer's instructions (MethylFlash, Epigentek, #P-1034). The positive DNA controls were prepared *via* serial dilutions (Table A.13).

Tube no.	Stock added (μL)	TE buffer added (μL)	Total volume (μL)	Final concentration ($\text{ng} \cdot \mu\text{l}^{-1}$)
1	5 μL of 10 $\text{ng}/\mu\text{L}$	20	25	2
2	12.5 μL of 1	12.5	25	1
3	10 μL of 2	15	25	0.4
4	12.5 μL of 3	12.5	25	0.2
5	12.5 μL of 4	12.5	25	0.1

Table A.13 Preparation of positive control dilutions for ELISA assay

All samples were added to the 48-well plate and the plate was sealed and incubated for 1 h at 37°C (Table A.13). After incubation, the wells were washed with buffer (3 x 150 μL) and a solution of capture antibody (1:1000, 50 μL) was added to each well. The plate was covered and incubated at room temperature for 1 h. The wells were subsequently washed (3 x 150 μL) and incubated with detection antibody (1:2000, 50 μL) at room temperature for 30 min. After removal of the detection antibody and

subsequent washes (4 x 150 μL), enhancer solution (1:5000, 50 μL) was added and the plate was incubated for room temperature for 30 min. After the final washes (5 x 150 μL), developer solution (100 μL) was added to each well and the plate was incubated for 5-10 min away from light. The solutions turned blue in the presence of sufficient methylated DNA. The final step was the addition of stop solution (100 μL). Final absorbance readings were measured at $\lambda = 450 \text{ nm}$ on a VarioSkan Flash Multimode Platerreader (Thermo Scientific).

	1	2	3	4
A	ME3 (-)	ME3 (-)	1.5 μM dAC	1.5 μM dAC
B	10 ng	10 ng	4.5 μM dAC	4.5 μM dAC
C	5 ng	5 ng	13.5 μM dAC	13.5 μM dAC
D	2 ng	2 ng	40.5 μM dAC	40.5 μM dAC
E	1 ng	1 ng	121.5 μM dAC	121.5 μM dAC
F	0.5 ng	0.5 ng	363.5 μM dAC	363.5 μM dAC
G	Media	Media	Media	Media
H	Media	Media	Media	Media

Table A.14 Assay setup for ELISA

	1	2	3	4
A	0.051	0.055	0.077	0.077
B	0.289	0.724	0.052	0.051
C	0.179	0.495	0.058	0.051
D	0.149	0.332	0.048	0.047
E	0.160	0.220	0.052	0.053
F	0.098	0.118	0.061	0.047
G	0.050	0.058	0.057	0.048
H	0.092	0.086	0.100	0.054

Table A.15 Absorbance readings from first ELISA trial

	1	2	3	4
A	0.037	0.036	0.044	0.044
B	0.125	0.250	0.048	0.058
C	0.081	0.096	0.038	0.039
D	0.053	0.067	0.035	0.035
E	0.048	0.054	0.036	0.038
F	0.047	0.049	0.035	0.038
G	0.042	0.038	0.039	0.039
H	0.060	0.091	0.037	0.045

Table A.16 Absorbance readings from second ELISA trial

A.4 Mass spectrometry analysis

A.4.1 Triple quadrupole mass spectrometer

The mass spectrometry analysis for the stability studies were performed on a Triple Quadrupole 6460 Mass Spectrometer (Agilent Technologies) fitted with an Infinity 1260 LC system (Agilent) and a Hypersil Gold C₁₈ column, 100 x 2.1 mm, 1.9 μm (Thermo Scientific, UK), using a gradient of water and acetonitrile with 0.1% formic acid (Table A.17). External calibration was performed and 2'-deoxycytidine (dC) was used as an internal standard.

Time (min)	Solvent A (%)	Solvent B (%)	Flow rate ($\mu\text{L}\cdot\text{min}^{-1}$)
0	98	2	200
20	2	98	200
25	98	2	200
27	98	2	200

Table A.17 LC-MS method conditions

A.4.2 MRM optimisation

MRM optimisation was carried out using the MassHunter Optimizer Software (Agilent) to determine the most suitable data acquisition parameters for each individual compound analysed. The software automated the selection of the best precursor ions, fragmentor voltages for each precursor ions, the best product ions and the collision energy values for each transition.

1 ng/ μL solutions of dAC, photocaged analogue (3'-CD, 5'-CD or 3'-CdC) and DEACM-OH were prepared in HPLC grade water and the samples were introduced to the system *via* an automatic injector. The fragmentor voltage range was set to 50-300 V with 5 step increments, the collision energy range was between 20 and 50 V and the cell accelerator voltage was 4V. Precursor ions were monitored with a 50 m/z low mass cutoff, in positive and negative ion mode. The optimisation duration for each analyte was around 3 h.

After the completion of the optimisation, the parameters with the highest product and precursor ion intensity were selected from Table A.18. The optimised conditions were incorporated into the LC method and were included in all subsequent LC-MRM methods.

Compound name	Precursor ion	Fragmentor voltage (V)	Product ion	Collision energy (V)	Measured abundance
3'-CD	502	105	292	20	40083
3'-CD	502	105	248	40	19867
3'-CD	502	105	112.9	32	4852
3'-CD	502	105	98.9	28	3812
5'-CD	502	110	80.9	50	35140
5'-CD	502	110	372	20	65214
5'-CD	502	110	230	28	19175
5'-CD	502	110	215	40	9476
3'-CdC	501	115	112	20	159129
3'-CdC	501	115	292.1	20	18152
3'-CdC	501	115	248.1	40	12193
3'-CdC	501	115	147.1	50	87
DEACM-OH	248	125	203.9	28	375175
DEACM-OH	248	125	176	40	257117
DEACM-OH	248	125	219	20	327139
DEACM-OH	248	125	162	36	60083
dAC	229	80	112.9	20	23515
dAC	229	80	112.2	20	13810
dAC	229	80	80.9	24	5538
dAC	229	80	69	44	4108
dC	228	55	111.9	20	248685
dC	228	55	94.9	44	54257
dC	228	55	69	40	21516
dC	228	55	137.1	20	50

Table A.18 MRM optimisation results

A.4.3 MRM LC-MS calibration of analytes

Calibration curves were constructed for each analyte in order to accurately quantify them in stability mixtures. Known concentrations (25 - 1000 $\text{pg}/\mu\text{L}$) of dAC, DEACM-OH and 3'-CD) were prepared and analysed using the optimised MRM LC methods and the integrals of the corresponding peaks were measured. The fitted calibration curves for dAC, DEACM-OH and 3'-CD are shown in Figure A.7.

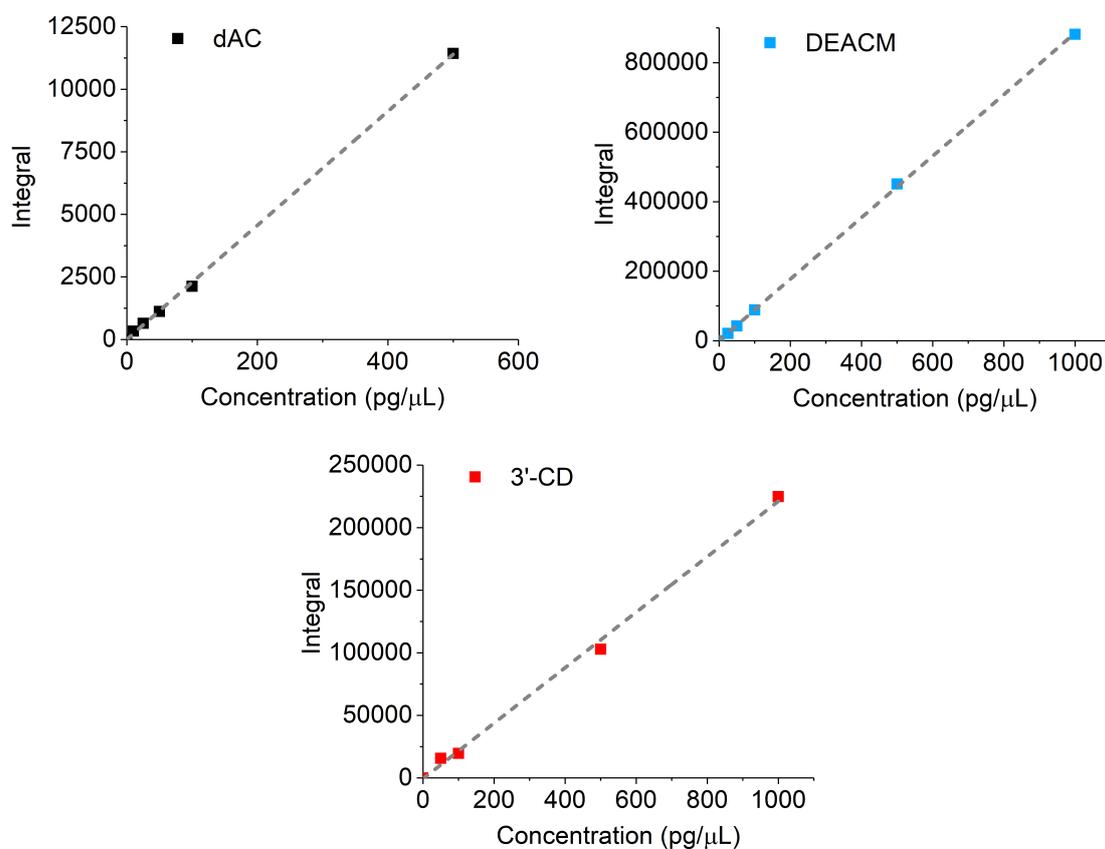


Fig. A.7 MRM LC-MS calibration of dAC, DEACM and 3'-CD

A.4.4 Sample preparation and stability experiments

All calibration solutions and stability mixtures were prepared from 10 mM stocks of the corresponding analytes (dAC, DEACM-OH or 3'-CD) in DMSO. A 5 ng/ μ L solution was made initially that was further diluted into the desired concentrations. The solutions were prepared using HPLC grade water or 100 mM HEPES buffer pH 7.2. All solutions were spiked with 500 pg/ μ L dC internal standard.

For all stability experiments, 750 pg/ μ L of the photocaged analogue or dAC was prepared (1 mL) and the solution was analysed at $t=0, 30, 60, 120, 180, 240, 300$ and 360 min *via* LC-MS/MS. The stability mixtures were either incubated at room temperature or 37 °C. For room temperature measurements, the solutions were left in the LC sample tray and an automatic sequence was set up to withdraw aliquots at the desired time points. Inbetween hourly injections, an additional blank containing water was also run to eliminate sample carry-over.

For stability studies at 37 °C, the mixtures were incubated in a ThermoMixer heating block (Eppendorf). At each time point, the entire solution was transferred to a MS vial for injection into the LCMS. After the injection, the solution was returned to the eppendorf tube and the incubation continued. Similar to room temperature studies, additional blanks containing water were included to eliminate sample carry-over. Each stability run was carried out in triplicates.

A.4.5 Quantification

Stability data was quantified using the MassHunter Workstation Software. After each time-point analysis (LC run), the peaks corresponding to the individual analytes (dAC, DEACM-OH, photocaged analogue and dC internal standard) were extracted from the main chromatogram and their areas were integrated. The integrals of the main analytes were then normalised against the integrals of the internal standard.

A.4.6 Summary of measured stability profiles

The measured stabilities and standard errors (SE) are summarised in Tables A.19

Time (min)	% dAC	SE	3'-CD	SE	5'-CD	SE	3'-CdC	SE
0	100	0.000	100	0.000	100	0.000	100	0.000
30	58	1.180	85	7.583	89	1.638	96	0.702
60	56	1.044	73	0.574	82	2.022	93	0.682
120	53	0.902	59	4.298	75	1.211	86	2.714
180	49	1.730	51	3.264	67	0.398	87	0.554
240	48	3.149	47	1.274	60	0.258	82	1.230
300	46	2.290	42	0.789	56	0.327	83	0.738
360	44	1.783	38	1.819	51	0.317	80	1.505

(a) Water, RT

Time (min)	dAC	SE	3'-CD	SE	3'-CdC	SE
0	100	0.000	100	0.000	100	0.000
30	87	2.964	49	7.534	68	1.274
60	83	2.501	39	5.037	62	2.022
120	80	4.948	33	0.478	55	4.037
180	75	3.867	27	3.670	49	1.505
240	71	5.309	25	1.277	44	1.277
300	67	3.294	20	0.216	40	3.264
360	64	3.524	15	0.252	37	0.789

(b) Water, 37 °C

Time (min)	3'-CD	SE	3'-CdC	SE	5'-CD	SE
0	100	0.000	100	0.000	100	0.000
30	94	0.895	99	0.895	95	1.577
60	90	0.838	98	0.838	92	2.019
120	83	0.978	98	0.978	89	1.723
180	78	0.355	95	0.355	85	0.704
240	73	1.299	92	0.706	83	2.034
300	67	2.358	89	0.834	82	1.481
360	64	0.969	88	0.969	79	1.842

(c) HEPES, RT

Time (min)	3'-CD	SE	3'-CdC	SE
0	100	0.000	100	0.000
30	67	1.667	88	1.577
60	58	0.706	80	2.019
120	48	2.470	76	1.723
180	40	2.240	70	0.704
240	34	1.392	61	2.034
300	27	0.815	55	1.481
360	22	1.199	53	1.842

(d) HEPES, 37 °C

Table A.19 Summary of LC-MS/MS measurements of stability studies

Appendix B

Publications

1. Ha Phuong Nguyen, Sabrina Stewart, Mikiembo N Kukwikila, Sioned Fon Jones, Daniel Offenbartl-Stiegert, Shiqing Mao, Shankar Balasubramanian, Stephan Beck, Stefan Howorka: **Opto-epigenetic modulation of DNA methylation with a photo-responsive small molecule approach**, *Angewandte Chemie International Edition* 10.1002/anie.201901139

Bibliography

- (1) S. Clancy and W. Brown, *Nat. Educ.*, 2008, **1**, 101.
- (2) J. M. Berg, J. L. Tymoczko and L. Stryer, in *Biochemistry*, ed. W. H. Freeman, New York, 5th, 2002, ch. 3.
- (3) N. Carey, *The Epigenetics Revolution*, Icon Books, 2012.
- (4) C. Deans and K. A. Maggert, *Genetics*, 2015, **199**, 887–896.
- (5) A. Bird, *Nature*, 2007, **447**, 396–398.
- (6) E. Li, *Nat. Rev. Genet.*, 2002, **3**, 662–73.
- (7) A. D. Goldberg, C. D. Allis and E. Bernstein, *Cell*, 2007, **128**, 635–638.
- (8) C. H. Waddington, *The Strategy of The Genes*, Allen & Unwin, 1957.
- (9) N. Gilbert and B. Ramsahoye, *Brief. Funct. Genomic Proteomic*, 2005, **4**, 129–142.
- (10) B. Li, M. Carey and J. L. Workman, *Cell*, 2007, **128**, 707–719.
- (11) M. A. Dawson, *Science*, 2017, **1**, 379–385.
- (12) B. E. Bernstein, A. Meissner and E. S. Lander, *Cell*, 2007, **128**, 669–681.
- (13) B. D. Strahl and C. D. Allis, *Nature*, 2000, **403**, 41–45.
- (14) D. Holoch and D. Moazed, *Nat. Rev. Genet.*, 2015, **16**, 10–11.
- (15) K. D. Robertson and P. A. Jones, *Carcinogenesis*, 2000, **21**, 461–467.
- (16) B. Jin, Y. Li and K. D. Robertson, *Genes Cancer*, 2011, **2**, 607–617.
- (17) J. S. Hagood, *Physiology*, 2014, **29**, 177–85.
- (18) A. Bird, *Genes Dev.*, 2002, **16**, 6–21.
- (19) R. J. Klose and A. P. Bird, *Trends Biochem. Sci.*, 2006, **31**, 89–97.
- (20) E. Prokhortchouk and P. A. Defossez, *Biochim. Biophys. Acta*, 2008, **1783**, 2167–2173.
- (21) R. S. Illingworth and A. P. Bird, *FEBS Lett.*, 2009, **583**, 1713–1720.
- (22) A. J. A. Sharp, E. Stathaki, E. Migliavacca *et al.*, *Genome Res.*, 2011, **21**, 1592–1600.
- (23) E. Li, C. Beard and R. Jaenisch, *Nature*, 1993, **366**, 362–365.
- (24) P. A. Jones, *Nat. Rev. Genet.*, 2012, **13**, 484–92.
- (25) R. Z. Jurkowska, T. P. Jurkowski and A. Jeltsch, *ChemBioChem*, 2011, **12**, 206–22.
- (26) T. Chen, Y. Ueda, J. E. Dodge, Z. Wang and E. Li, *Mol. Cell. Biol.*, 2003, **23**, 5594–605.
- (27) M. Okano, D. W. Bell, D. A. Haber and E. Li, *Cell*, 1999, **99**, 247–257.
- (28) J. Song, O. Rechko, T. H. Bestor and D. J. Patel, *Science*, 2011, **331**, 1036–40.
- (29) M. Bostick, J. K. Kim, P.-O. Esteve, A. Clark, S. Pradhan and S. E. Jacobsen, *Science*, 2007, **317**, 1760–1765.

- (30) A. Hermann, R. Goyal and A. Jeltsch, *J. Biol. Chem.*, 2004, **279**, 48350–48359.
- (31) Z. M. Svedruzić, *Curr. Med. Chem.*, 2008, **15**, 92–106.
- (32) L. D. Moore, T. Le and G. Fan, *Neuropsychopharmacology*, 2012, **38**, 23–38.
- (33) S. Kriaucionis and N. Heintz, *Science*, 2009, **324**, 929–30.
- (34) M. Tahiliani, K. P. Koh, Y. Shen *et al.*, *Science*, 2009, **324**, 930–5.
- (35) R. M. Kohli and Y. Zhang, *Nature*, 2013, **502**, 472–9.
- (36) S. Ito, L. Shen, Q. Dai *et al.*, *Science*, 2011, **333**, 1300–3.
- (37) W. a. Pastor, L Aravind and A. Rao, *Nat. Rev. Mol. Cell Biol.*, 2013, **14**, 341–56.
- (38) M. W. Kellinger, C. X. Song, J Chong, X. Y. Lu, C He and D Wang, *Nat Struct Mol Biol*, 2012, **19**, 831–833.
- (39) M. Iurlaro, G. R. McInroy, H. E. Burgess *et al.*, *Genome Biol.*, 2016, **17**, 141.
- (40) J. Song and G. P. Pfeifer, *BioEssays*, 2016, **38**, 1038–1047.
- (41) B. Sudhamalla, D. Dey, M. Breski and K. Islam, *Anal. Biochem.*, 2017, **534**, 28–35.
- (42) H. Wu, A. C. D'Alessio, S. Ito *et al.*, *Nature*, 2011, **473**, 389–93.
- (43) D. D. De Carvalho, J. S. You and P. a. Jones, *Trends Cell Biol.*, 2010, **20**, 609–17.
- (44) F. M. Piccolo, H. Bagci, K. E. Brown *et al.*, *Mol. Cell*, 2013, **49**, 1023–1033.
- (45) M. Ko, Y. Huang, A. M. Jankowska *et al.*, *Nature*, 2010, **468**, 839–43.
- (46) T. P. Gu, F Guo, H Yang *et al.*, *Nature*, 2011, **477**, 606–610.
- (47) Z. D. Smith and A. Meissner, *Nat. Rev. Genet.*, 2013, **14**, 204–20.
- (48) B. C. Richardson, *J. Nutr.*, 2002, 2401–2405.
- (49) R Rizwana and P. J. Hahn, *J. Cell Sci.*, 1999, **112**, 4513–9.
- (50) H. Meng, H. Meng, Y. Cao *et al.*, *Int. J. Biol. Sci.*, 2015, **11**, 604–617.
- (51) K. D. Robertson, *Nat. Rev. Genet.*, 2005, **6**, 597–610.
- (52) R. D. Nicholls, S Saitoh and B Horsthemke, *Trends Genet.*, 1998, **14**, 194–200.
- (53) S. Choufani, C. Shuman and R. Weksberg, *Am. J. Med. Genet. Part C Semin. Med. Genet.*, 2010, **154**, 343–354.
- (54) G. Egger, G. Liang, A. Aparicio and P. A. Jones, *Nature*, 2004, **429**, 457–463.
- (55) M. Esteller, *Annu. Rev. Pharmacol. Toxicol.*, 2005, **45**, 629–656.
- (56) M. Ehrlich, *Clin. Immunol.*, 2003, **109**, 17–28.
- (57) M Ehrlich, C Sanchez, C Shao *et al.*, *Autoimmunity*, 2008, **41**, 253–271.
- (58) R. E. Amir, I. B. Van den Veyver, M. Wan, C. Q. Tran, U. Francke and H. Y. Zoghbi, *Nat. Genet.*, 1999, **23**, 185–188.
- (59) P. A. Jones, *Oncogene*, 2002, **21**, 5358–60.
- (60) A. P. Feinberg and B. Tycko, *Nat. Rev. Cancer*, 2004, **4**, 143–153.
- (61) P. a. Jones and S. B. Baylin, *Cell*, 2007, **128**, 683–692.
- (62) M. Varela-Rey, A. Woodhoo, M.-L. Martinez-Chantar, J. M. Mato and S. C. Lu, *Alcohol Res.*, 2013, **35**, 25–35.
- (63) A. P. Feinberg and B. Vogelstein, *Nature*, 1983, **301**, 89–92.
- (64) M. A. Gama-Sosa, V. A. Slagel, R. W. Trewyn *et al.*, *Nucleic Acids Res.*, 1983, **1**, 6883–6894.
- (65) M. Ehrlich, *Epigenomics*, 2009, **1**, 239–59.
- (66) M. Esteller, *Oncogene*, 2002, **21**, 5427–5440.
- (67) J. G. Herman and S. B. Baylin, *N. Engl. J. Med.*, 2003, **349**, 2042–54.
- (68) S. B. Baylin, *Nat. Clin. Pract. Oncol.*, 2005, **2**, S4–S11.

- (69) G. Gopisetty, K. Ramachandran and R. Singal, *Mol. Immunol.*, 2006, **43**, 1729–1740.
- (70) P. A. Jones, J. Issa and S. Baylin, *Nat. Rev. Genet.*, 2016, **17**, 630–41.
- (71) J. D. Tompkins, C. Hall, V. C.-y. Chen, A. L. Xuejun, X. Wu and D. Hsu, *Proc. Natl. Acad. Sci.*, 2012, **109**, 12544–12549.
- (72) P. Voigt and D. Reinberg, *Nat. Biotechnol.*, 2013, **31**, 1097–9.
- (73) T. P. Jurkowski, M. Ravichandran and P. Stepper, *Clin. Epigenetics*, 2015, **7**, 18.
- (74) L. Yu, J. Batara and B. Lu, in *Mod. Tools Genet. Eng.* 2016, pp. 163–186.
- (75) M. L. De Groote, P. J. Verschure and M. G. Rots, *Nucleic Acids Res.*, 2012, **40**, 10596–10613.
- (76) G. Kubik and D. Summerer, *ChemBioChem*, 2016, **17**, 975–980.
- (77) M. L. Maeder, J. F. Angstman, M. E. Richardson *et al.*, *Nat. Biotechnol.*, 2013, **31**, 1137–42.
- (78) E. M. Mendenhall, K. E. Williamson, D. Reyon *et al.*, *Nat. Biotechnol.*, 2013, **31**, 1133–6.
- (79) S. Konermann, M. D. Brigham, A. E. Trevino *et al.*, *Nature*, 2013, **500**, 472–6.
- (80) D. Carroll, *Nat. Biotechnol.*, 2013, **31**, 807–809.
- (81) N. Rusk, *Nat. Methods*, 2014, **11**, 28–28.
- (82) E. Pennisi, *Science*, 2013, **341**, 833–836.
- (83) J. I. McDonald, H. Celik, L. E. Rois *et al.*, *Biol. Open*, 2016, **5**, 866–74.
- (84) X. S. Liu, H. Wu, X. Ji *et al.*, *Cell*, 2016, **167**, 233–247.
- (85) A. Vojta, P. Dobrinic, V. Tadic *et al.*, *Nucleic Acids Res.*, 2016, **44**, 5615–5628.
- (86) H. Ma, N. Marti-Gutierrez, S.-W. Park *et al.*, *Nature*, 2017, **548**, 413–419.
- (87) J. M. Foulks, K. M. Parnell, R. N. Nix *et al.*, *J. Biomol. Screen.*, 2012, **17**, 2–17.
- (88) B. Segura-Pacheco, C. Trejo-Becerril, E. Perez-Cardenas *et al.*, *Clin. Cancer Res.*, 2003, **9**, 1596–1603.
- (89) M. Z. Fang, Y. Wang, N. Ai *et al.*, *Cancer Res*, 2003, **63**, 7563–7570.
- (90) P. Siedlecki, R. G. Boy, T. Musch *et al.*, *J. Med. Chem.*, 2006, **49**, 678–683.
- (91) S. H. Song, S. W. Han and Y. J. Bang, *Drugs*, 2011, **71**, 2391–2403.
- (92) T. Suzuki, R. Tanaka, S. Hamada, H. Nakagawa and N. Miyata, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 1124–1127.
- (93) R. J. Amato, *Clin. Genitourin. Cancer*, 2007, **5**, 422–426.
- (94) R. J. Amato, J. Stephenson, S. Hotte *et al.*, *Cancer Invest.*, 2012, **30**, 415–421.
- (95) P. M. Howell, Z. Liu and H. T. Khong, *Pharmaceuticals*, 2010, **3**, 2022–2044.
- (96) F. Šorm, A. Pískala, A. Čihák and J. Veselý, *Experientia*, 1964, **20**, 202–203.
- (97) L. H. Li, E. J. Olin, H. H. Buskirk and L. M. Reineke, *Cancer Res.*, 1970, **30**, 2760–2769.
- (98) D. D. Von Hoff, M Slavik and F. M. Muggia, *Ann. Intern. Med.*, 1976, **85**, 237–245.
- (99) P. A. Jones and S. M. Taylor, *Cell*, 1980, **20**, 85–93.
- (100) D. de Vos and W. van Overveld, *Ann. Hematol. Suppl.*, 2005, **84**, 3–8.
- (101) S. D. Gore, C. Jones and P. Kirkpatrick, *Nat. Rev. Drug Discov.*, 2006, **5**, 891–892.
- (102) E. H. Estey, *Leukemia*, 2013, **27**, 1803–12.
- (103) X. Yang, F. Lay, H. Han and P. A. Jones, *Trends Pharmacol. Sci.*, 2010, **31**, 536–546.

- (104) V. L. Damaraju, D. Mowles, S. Yao *et al.*, *Nucleosides Nucleotides Nucleic Acids*, 2012, **31**, 236–255.
- (105) M. Rius, C. Stresemann, D. Keller *et al.*, *Mol. Cancer Ther.*, 2009, **8**, 225–31.
- (106) K. Ueda, M. Hosokawa and S. Iwakawa, *Biol. Pharm. Bull.*, 2015, **38**, 1113–1119.
- (107) R. L. Momparler and D. Derse, *Biochem. Pharmacol.*, 1979, **28**, 1443–4.
- (108) R. L. Momparler, *Semin. Hematol.*, 2005, **42**, S9–S16.
- (109) J. Bouchard and R. L. Momparler, *Mol. Pharmacol.*, 1983, **24**, 109–114.
- (110) J. Aimiuwu, H. Wang, P. Chen *et al.*, *Blood*, 2012, **119**, 5229–5238.
- (111) C. Stresemann and F. Lyko, *Int. J. Cancer*, 2008, **123**, 8–13.
- (112) R. L. Momparler, J. Samson and L. F. Momparler, *Cancer Chemother. Pharmacol.*, 1984, **13**, 191–194.
- (113) S. Öz, G. Raddatz, M. Rius *et al.*, *Nucleic Acids Res.*, 2014, **42**, e152.
- (114) K. Patel, J. Dickson, S. Din, K. Macleod, D. Jodrell and B. Ramsahoye, *Nucleic Acids Res.*, 2010, **38**, 4313–4324.
- (115) D. V. Santi, a Norment and C. E. Garrett, *Proc. Natl. Acad. Sci.*, 1984, **81**, 6993–6997.
- (116) L. Pleyer and R. Greil, *Drug Metab. Rev.*, 2015, **47**, 252–79.
- (117) K. Ghoshal, J. Datta, S. Majumder *et al.*, *Mol. Cell. Biol.*, 2005, **25**, 4727–4741.
- (118) J. Datta, K. Ghoshal, T. Motiwala and S. T. Jacob, *Genes Cancer*, 2012, **3**, 71–81.
- (119) M. Schaefer, S. Hagemann, K. Hanna and F. Lyko, *Cancer Res*, 2009, 8127–8133.
- (120) J. Diesch, A. Zwick, A.-K. Garz, A. Palau, M. Buschbeck and K. S. Götze, *Clin. Epigenetics*, 2016, **8**, 71.
- (121) T. Qin, J. Jelinek, J. Si, J. Shu and J.-P.J.-P. J. Issa, *Blood*, 2009, **113**, 659–667.
- (122) C. Arimany-Nardi, E. Errasti-Murugarren, G. Minuesa *et al.*, *Br. J. Pharmacol.*, 2014, **171**, 3868–3880.
- (123) N. Eliopoulos, D. Cournoyer and R. L. Momparler, *Cancer Chemother. Pharmacol.*, 1998, **42**, 373–378.
- (124) D. Lavelle, K. Vaitkus, Y. Ling *et al.*, *Blood*, 2012, **119**, 1240–1247.
- (125) Q. Ebrahim, R. Z. Mahfouz, K. P. Ng and Y. Saunthararajah, *Oncotarget*, 2012, **3**, 1137–45.
- (126) M. Hajek, I. Votruba, A. Holy, M. Krecmerova and E. Tloustova, *Biochem. Pharmacol.*, 2008, **75**, 965–972.
- (127) J. Vesely and A. Piskala, *Cancer Res.*, 1984, **44**, 5165–5168.
- (128) M. Fojtova, A. Piskala, I. Votruba, M. Otmar, E. Bartova and A. Kovarik, *Pharmacol. Res.*, 2007, **55**, 16–22.
- (129) M. J. Davies, P. R. Jenkins, L. J. S. Prouse, D. R. Russell and D. de Vos, *Acta Crystallogr. C*, 1991, **47**, 1418–1420.
- (130) S. C. Navada, J. Steinmann, M. Lübbert and L. R. Silverman, *J. Clin. Invest.*, 2014, **124**, 40–46.
- (131) H. M. Kantarjian and J. P. Issa, *Semin. Hematol.*, 2005, **42**, S17–S22.
- (132) H. Wu and Y. Zhang, *Cell*, 2014, **156**, 45–68.
- (133) O. I. Kulaeva, S. Draghici, L. Tang, J. M. Kraniak, S. J. Land and M. A. Tainsky, *Oncogene*, 2003, **22**, 4118–27.
- (134) L. J. Rush, A. Raval, P. Funchain *et al.*, *Cancer Res.*, 2004, **64**, 2424–2433.
- (135) D. Gius, C. M. Bradbury, L. Sun *et al.*, *Cancer*, 2005, **104**, 1789–1793.

- (136) K Schmelz, N Sattler, M Wagner, M Lübbert, B Dörken and I Tamm, *Leukemia*, 2005, **19**, 103–111.
- (137) D. S. Shames, L. Girard, B. Gao *et al.*, *PLoS Med.*, 2006, **3**, 2244–2263.
- (138) I. Ibanez De Caceres, E. Dulaimi, A. M. Hoffman, T. Al-Saleem, R. G. Uzzo and P. Cairns, *Cancer Res.*, 2006, **66**, 5021–5028.
- (139) G. Heller, W. M. Schmidt, B. Ziegler *et al.*, *Cancer Res.*, 2008, **68**, 44–54.
- (140) F. Haller, J. D. Zhang, E. A. Moskalev *et al.*, *Int. J. Cancer*, 2015, **136**, 1013–1023.
- (141) M. F. Paz, M. F. Fraga, S. Avila *et al.*, *Cancer Res.*, 2003, **63**, 1114–1121.
- (142) K. Ogoshi, S.-i. Hashimoto, Y. Nakatani *et al.*, *Genomics*, 2011, **98**, 280–287.
- (143) K. E. Varley, J. Gertz, K. M. Bowling *et al.*, *Genome Res.*, 2013, **23**, 555–567.
- (144) S. Yamashita, Y. Tsujino, K. Moriguchi, M. Tatematsu and T. Ushijima, *Cancer Sci.*, 2006, **97**, 64–71.
- (145) D. Mossman, K.-T. Kim and R. J. Scott, *BMC Cancer*, 2010, **10**, 366.
- (146) M.-P. Ramos, N. A. Wijetunga, A. S. McLellan, M. Suzuki and J. M. Greally, *Epigenetics Chromatin*, 2015, **8**, DOI: 10.1186/s13072-015-0004-x.
- (147) H. Araki, K. Yoshinaga, P. Boccuni, Y. Zhao, R. Hoffman and N. Mahmud, *Blood*, 2007, **109**, 3570–3578.
- (148) C. Schmittwolf, N. Kirchhof, A. Jauch *et al.*, *EMBO J.*, 2005, **24**, 554–566.
- (149) B. S. Pace, L. Liu, B. Li and L. H. Makala, *Exp. Biol. Med.*, 2015, **240**, 1050–1064.
- (150) D. Biswas and P. Jiang, *Int. J. Mol. Sci.*, 2016, **17**, 226–239.
- (151) E. Jabbour, J. P. Issa, G. Garcia-Manero and H. Kantarjian, *Cancer*, 2008, **112**, 2341–2351.
- (152) J. Nie, L. Liu, X. Li and W. Han, *Cancer Lett.*, 2014, **354**, 12–20.
- (153) X. Li, Q. Mei, J. Nie, X. Fu and W. Han, *Expert Rev. Clin. Immunol.*, 2015, **11**, 363–75.
- (154) J. A. Beisler, *J. Med. Chem.*, 1978, **21**, 204–208.
- (155) D. K. Rogstad, J. L. Herring, J. A. Theruvathu *et al.*, *Chem. Res. Toxicol.*, 2009, **22**, 1194–1204.
- (156) B. Chabner, J. Drake and D. Johns, *Biochem. Pharmacol.*, 1973, **22**, 2763–2765.
- (157) J. P. J. Issa and H. M. Kantarjian, *Clin. Cancer Res.*, 2009, **15**, 3938–3946.
- (158) N. J. Yogelzang, J. E. Herndon, C. Cirrincione *et al.*, *Cancer*, 1997, **79**, 2237–42.
- (159) M. Matoušová, I. Votruba, M. Otmar, E. Tloušťová, J. Günterová and H. Mertlíková-Kaiserová, *Epigenetics*, 2011, **6**, 769–776.
- (160) M. Krečmerová and M. Otmar, *Future Med. Chem.*, 2012, **4**, 991–1005.
- (161) J. I. Mullins, L. Heath, J. P. Hughes *et al.*, *PLoS One*, 2011, **6**, 15135.
- (162) J. L. Holleran, R. A. Parise, E. Joseph *et al.*, *Cancer Ther. Preclin.*, 2005, **11**, 3862–3869.
- (163) J. H. Beumer, J. L. Eiseman, R. A. Parise *et al.*, *Clin. Cancer Res.*, 2006, **12**, 7483–7491.
- (164) E. M. Newman and D. V. Santit, *Proc. Natl. Acad. Sci.*, 1982, **79**, 6419–6423.
- (165) R. Güimil García, A. S. Brank, J. K. Christman, V. E. Marquez and R. Eritja, *Antisense Nucleic Acid Drug Dev.*, 2001, **11**, 369–78.
- (166) H.-M. Byun, S. H. Choi, P. W. Laird *et al.*, *Cancer Lett.*, 2008, **266**, 238–248.
- (167) J.-P. J. Issa, G. Roboz, D. Rizzieri *et al.*, *Lancet. Oncol.*, 2015, **16**, 1099–1110.
- (168) J. Linnekamp, R. Butter, R. Spijker, J. Medema and H. van Laarhoven, *Cancer Treat. Rev.*, 2017, **54**, 10–23.

- (169) B. Brueckner, M. Rius, M. R. Markelova *et al.*, *Mol. Cancer Ther.*, 2010, **9**, 1256–64.
- (170) A. Ziemba, E. Hayes, B. B. Freeman, T. Ye and G. Pizzorno, *Chemother. Res. Pract.*, 2011, **2011**, 965826.
- (171) E. Laille, M. R. Savona, B. L. Scott, T. E. Boyd, Q. Dong and B. Skikne, *J. Clin. Pharmacol.*, 2014, **54**, 630–639.
- (172) D Lavelle, J Chin, K Vaitkus *et al.*, *Am. J. Hematol.*, 2007, **82**, 981–985.
- (173) Y. Zhang, J. Sun, Y. Gao *et al.*, *Mol. Pharm.*, 2013, **10**, 3195–202.
- (174) E. A. Griffiths, G Choy, S Redkar, P Taverna, M Azab and A. R. Karpf, *Drugs Future*, 2013, **38**, 535–543.
- (175) G. Mayer and A. Heckel, *Angew. Chem. Int. Ed. Engl.*, 2006, **45**, 4900–21.
- (176) S. A. Reis, B. Ghosh, J. A. Hendricks *et al.*, *Nat. Chem. Biol.*, 2016, **12**, 317–23.
- (177) L. R. Polstein and C. A. Gersbach, *Nat. Chem. Biol.*, 2015, **11**, 198–201.
- (178) Y. Nihongaki, S. Yamamoto, F. Kawano, H. Suzuki and M. Sato, *Chem. Biol.*, 2015, **22**, 169–174.
- (179) J. Hemphill, E. K. Borchardt, K. Brown, A. Asokan and A. Deiters, *J. Am. Chem. Soc.*, 2015, **137**, 5642–5645.
- (180) Y. Nihongaki, F. Kawano, T. Nakajima and M. Sato, *Nat. Biotechnol.*, 2015, **33**, 755–762.
- (181) P. Klán, T. Šolomek, C. G. Bochet *et al.*, *Chem. Rev.*, 2013, **113**, 119–91.
- (182) G. C. R. Ellis-Davies, *Nat. Methods*, 2007, **4**, 619–628.
- (183) E. W. Yankee and D. J. Cram, *J. Am. Chem. Soc.*, 1970, **92**, 6333–6335.
- (184) T. Solomek, S. Mercier, T. Bally and C. G. Bochet, *Photochem. Photobiol. Sci.*, 2012, **11**, 548–55.
- (185) E Reichmanis, B. C. Smith and R Gooden, *J. Polym. Sci. Polym. Chem. Ed.*, 1985, **23**, 1–8.
- (186) A. P. Pelliccioli and J. Wirz, *Photochem. Photobiol. Sci.*, 2002, **1**, 441–458.
- (187) A Hasan, K. Stengele, H Giegrich *et al.*, *Tetrahedron*, 1997, **53**, 4247–4264.
- (188) M Beier and J. D. Hoheisel, *Nucleic Acids Res.*, 2000, **28**, E11.
- (189) H. Lusic, D. D. Young, M. O. Lively and A. Deiters, *Org. Lett.*, 2007, **9**, 1903–1906.
- (190) R. S. Givens and B. Matuszewski, *J. Am. Chem. Soc.*, 1984, **106**, 6860–6861.
- (191) A. Z. Suzuki, T. Watanabe, M. Kawamoto *et al.*, *Org. Lett.*, 2003, 2723–2726.
- (192) V. Hagen, S. Frings, B. Wiesner, S. Helm, U. B. Kaupp and J. Bendig, *ChemBioChem*, 2003, **4**, 434–42.
- (193) C. Bao, G. Fan, Q. Lin *et al.*, *Org. Lett.*, 2012, **14**, 572–5.
- (194) L. Fournier, C. Gauron, L. Xu *et al.*, *ACS Chem. Biol.*, 2013, **8**, 1528–1536.
- (195) M. Maheswara, V. Siddaiah, G. L. V. Damu, Y. K. Rao and C. V. Rao, *J. Mol. Catal. A Chem.*, 2006, **255**, 49–52.
- (196) M. K. Potdar, S. S. Mohile and M. M. Salunkhe, *Tetrahedron Lett.*, 2001, **42**, 9285–9287.
- (197) A. V. Pinheiro, P. Baptista and J. C. Lima, *Nucleic Acids Res.*, 2008, **36**, e90.
- (198) P. Bourbon, Q. Peng, G. Ferraudi, C. Stauffacher, O. Wiest and P. Helquist, *Bioorg. Med. Chem. Lett.*, 2013, **23**, 2162–5.
- (199) J. H. Kaplan, B Forbush and J. F. Hoffman, *Biochemistry*, 1978, **17**, 1929–1935.
- (200) V. Hagen, S. Frings, J Bendig, D. Lorenz, B. Wiesner and U. B. Kaupp, *Angew. Chem. Int. Ed.*, 2002, **41**, 3625–3628.

- (201) D. Geissler, W. Kresse, B. Wiesner, J. Bendig, H. Kettenmann and V. Hagen, *ChemBioChem*, 2003, **4**, 162–70.
- (202) T. Eckardt, V. Hagen, R. Schmidt and C. Schweitzer, *J. Am. Chem. Soc.*, 2002, **67**, 703–710.
- (203) S. Bühler, I. Lagoja, H. Giegrich, K.-P. Stengele and W. Pfeleiderer, *Helv. Chim. Acta*, 2004, **87**, 620–659.
- (204) X. Tang, J. Zhang, J. Sun, Y. Wang, J. Wu and L. Zhang, *Org. Biomol. Chem.*, 2013, **11**, 7814–7824.
- (205) J. Wu, S. Zhang, Q. Meng *et al.*, *Proc. Natl. Acad. Sci.*, 2007, **104**, 16462–16467.
- (206) S. G. Chaulk and A. M. MacMillan, *Nucleic Acids Res.*, 1998, **26**, 3173–3178.
- (207) A. S. Mathews, H. Yang, C. Montemagno *et al.*, *Org. Biomol. Chem*, 2016, **14**, 8278–8288.
- (208) J. Lu, S. C. Koo, N.-S. Li and J. A. Piccirilli, *Nucleosides Nucleotides Nucleic Acids*, 2015, **34**, 114–129.
- (209) A. Deiters, *Curr. Opin. Chem. Biol.*, 2009, **13**, 678–686.
- (210) Q. Liu and A. Deiters, *Acc. Chem. Res.*, 2014, **47**, 45–55.
- (211) L. Kröck and A. Hechel, *Angew. Chem. Int. Ed.*, 2005, **44**, 471–473.
- (212) G. Mayer, L. Kröck, V. Mikat, M. Engeser and A. Heckel, *ChemBioChem*, 2005, **6**, 1966–1970.
- (213) R. Ting, L. Lermer and D. M. Perrin, *J. Am. Chem. Soc.*, 2004, **126**, 12720–12721.
- (214) X. Tang and I. J. Dmochowski, *Org. Lett.*, 2005, **7**, 279–282.
- (215) T. Furuta, T. Watanabe, S. Tanabe, J. Sakyo and C. Matsuba, *Org. Lett.*, 2007, **9**, 4717–4720.
- (216) K. A. Scheidt, H. Chen, B. C. Follows, S. R. Chemler and D. S. Coffey, *J. Org. Chem.*, 1998, **3263**, 6436–6437.
- (217) Y. A. Kim, D. M. C. Ramirez, W. J. Costain, L. J. Johnston and R. Bittman, *Chem. Commun.*, 2011, **47**, 9236–8.
- (218) R. O. Schönleber, J. Bendig, V. Hagen and B. Giese, *Bioorg. Med. Chem.*, 2002, **10**, 97–101.
- (219) A. Nadler, G. Reither, S. Feng *et al.*, *Angew. Chem. Int. Ed.*, 2013, **52**, 6330–4.
- (220) M. André, S. Tarrit, M.-J. Couret *et al.*, *Org. Biomol. Chem.*, 2013, **11**, 6372–84.
- (221) T. Furuta, *J. Syn. Org. Chem. Jpn*, 2012, **70**, 1164–1169.
- (222) J. T. Weiss, J. C. Dawson, C. Fraser *et al.*, *J. Med. Chem.*, 2014, **57**, 5395–5404.
- (223) C. Baldoli, C. Rigamonti, S. Maiorana, E. Licandro, L. Falciola and P. R. Mussini, *Chem. Eur. J*, 2006, **12**, 4091–4100.
- (224) R. Ramapanicker, N. B. R. Baig, K. De and S. Chandrasekaran, *J. Pept. Sci.*, 2009, **15**, 849–855.
- (225) A. T. Khan and E. Mondal, *Synlett*, 2003, 694–698.
- (226) W. Reik, *Nature*, 2007, **447**, 425–432.
- (227) Y. Kobayashi and D. Absher, *Genome Res.*, 2011, **21**, 1017–1027.
- (228) M. P. Hanley, M. A. Hahn, A. X. Li *et al.*, *Oncogene*, 2017, 5035–5044.
- (229) M. Umer and Z. Herceg, *Antioxid. Redox Signal.*, 2013, **18**, 1972–1986.
- (230) E. J. Oakeley, *Pharmacol. Ther.*, 1999, **84**, 389–400.
- (231) P. W. Laird, *Nat. Rev. Genet.*, 2010, **11**, 191–203.
- (232) G. E. Lind, *Epigenomics*, 2017, **9**, 933–935.
- (233) K. C. Kuo, R. A. McCune, C. W. Gehrke, R. Midgett and M. Ehrlich, *Nucleic Acids Res.*, 1980, **8**, 4763–76.

- (234) C. Dahl and P. Guldberg, *Biogerontology*, 2003, **4**, 233–250.
- (235) M. F. Fraga, E. Uriol, L. Borja Diego *et al.*, *Electrophoresis*, 2002, **23**, 1677–1681.
- (236) M. F. Fraga, R. Rodriguez and M. J. Canal, *Electrophoresis*, 2000, **21**, 2990–2994.
- (237) M. Karimi, S. Johansson and T. J. Ekström, *Epigenetics*, 2006, **1**, 45–48.
- (238) S. Kurdyukov and M. Bullock, *Biology*, 2016, **5**, 3.
- (239) K. Lund, J. J. Cole, N. D. VanderKraats *et al.*, *Genome Biol.*, 2014, **15**, 406.
- (240) M. Y. So, Z. Tian, Y. S. Phoon *et al.*, *PLoS One*, 2014, **9**, DOI: 10.1371/journal.pone.0096404.
- (241) D. Kremer, S. Metzger, V. Kolb-bachofen and D. Kremer, *Anal. Biochem.*, 2012, **422**, 74–78.
- (242) L. Song, S. R. James, L. Kazim and A. R. Karpf, *Anal. Chem.*, 2005, **77**, 504–510.
- (243) S. Hagemann, O. Heil, F. Lyko and B. Brueckner, *PLoS One*, 2011, **6**, DOI: 10.1371/journal.pone.0017388.
- (244) Y. Oki, E. Aoki and J. P. J. Issa, *Crit. Rev. Oncol. Hematol.*, 2007, **61**, 140–152.
- (245) M. B. Omary, M. Mccaffrey and M. F. Kagnoff, *J. Cell. Biochem.*, 1992, **48**, 316–323.
- (246) D. K. Oelschlager, L. I. Talley, M. N. Barnes and E. E. Partridge, *Biotech. Histochem.*, 2003, **78**, 17–21.
- (247) H. Lee and J. U.N. B. Park, *Biomed. Reports*, 2017, 291–296.
- (248) X. Du, D. Lu, E. D. Daharsh, A. Yao, R. Dewoody and J.-a. Yao, *J. Pharmacol. Toxicol. Methods*, 2006, **54**, 164–172.
- (249) Y. H. Yang, H. Aloysius, D. Inoyama, Y. Chen and L. Q. Hu, *Acta Pharm. Sin. B*, 2011, **1**, 143–159.
- (250) T. L. Huang, A. Szekacs, T. Uematsu, E. Kuwano, A. Parkinson and B. D. Hammock, *Pharm. Res.*, 1993, **10**, 639–648.
- (251) T. Shi, D. Su, T. Liu *et al.*, *Proteomics*, 2012, **12**, 1074–1092.
- (252) C. Stresemann, B. Brueckner, T. Musch, H. Stopper and F. Lyko, *Cancer Res.*, 2006, **66**, 2794–800.
- (253) N. Zekri, R. F. Alamdari and A. Khalafi-Nezhad, *Bull Chem Soc Ethiop*, 2010, **24**, 299–304.
- (254) C. Mcguigan, M. Serpi, M. Slusarczyk *et al.*, *ChemistryOpen*, 2016, **5**, 227–235.
- (255) P. A.-a. Drugs, C.-h. Kim, V. E. Marquez, S. Broder, H. Mitsuya and J. S. Driscoll, *J. Med. Chem.*, 1987, 862–866.
- (256) K. Agrawal, V. Das, P. Vyas and M. Hajdúch, *Pharmacol. Ther.*, 2018, **188**, 45–79.
- (257) S. Bohacova, Z. Vanikova, L. P. Slavetinska and M. Hocek, *Org. Biomol. Chem.*, 2018, **16**, DOI: 10.1039/c8ob01106k.
- (258) S. Bohacova, L. Ludvikova, L. P. Slavetinska, Z. Vanikova, P. Klan and M. Hocek, *Org. Biomol. Chem.*, 2018, **16**, 1527–1535.
- (259) M. Noguchi, M. Skwarczynski, H. Prakash *et al.*, *Bioorg. Med. Chem.*, 2008, **16**, 5389–97.
- (260) G. I. Peterson, M. B. Larsen and A. J. Boydston, *Macromolecules*, 2012, **45**, 7317–7328.
- (261) M. A. Dewit and E. R. Gillies, *J. Am. Chem. Soc.*, 2009, 18327–18334.
- (262) P. W. Hollenbach, A. N. Nguyen, H. Brady *et al.*, *PLoS One*, 2010, **5**, e9001.

-
- (263) A. Hassner, D. Yagudayev, T. K. Pradhan, A. Nudelman and B. Amit, *Org. Lett.*, 2007, 2405–2409.
- (264) C. Sotelo-vazquez, R. Quesada-cabrera, M. Ling *et al.*, *Adv. Funct. Mater.*, 2017, **27**, 1–10.
- (265) H. Joo, *Korean J Chem Eng*, 2006, **23**, 931–934.
- (266) C. Swenton, J. S., Bonke, B. R., Chen, C. P., Chou, *J. Org. Chem.*, 1989, **2**, 51–58.
- (267) B. E. Watkins, J. S. Kiely and H. Rapoport, *J. Am. Chem. Soc.*, 1982, **104**, 5702–5708.

